Pennsylvania 2020 Voting Analysis Report

11-16-20

Table of Contents

Executive Summary

This scientific analysis of the reported Pennsylvania (PA) 2020 Presidential voting results, is a non-partisan effort by unpaid citizens and volunteer experts. Our only objective is to play a small roll in helping assure that all legal PA votes are counted, *and* that only legal PA votes are counted.

Whether Donald Trump or Joseph Biden wins is not of concern in this analysis — the scientists involved with the report just want the election results to truly reflect the wishes of Pennsylvania voting citizens.

Since there are multiple reports of voting chicanery circulating the Internet, a collection of statisticians and other scientists volunteered to examine the reported PA results from a scientific statistical perspective.

We feel that the best way to do this is to start by putting ourselves in the shoes of bad actors — and then considering how they might go about changing the wishes of PA citizens, into a different result. Some of the actions they might take are:

1 - Keep ineligible people (e.g. deceased, moved, etc.) on the voting roles.
   (This would disguise actual voter participation rates, allow fabricated votes to be submitted in their names, etc.)
2 - Get legislation passed that did not require in-person voter identification.
   (This would make it easier for non-citizens, felons, etc. to vote.)
3 - Encourage a much higher percentage of voting by mail.
   (This would make it much easier to manipulate, as in-person checking is a more secure way to keep track of actual registered citizens, etc.)
4 - Discard envelopes and other identifying materials from mail-in votes.
   (This makes it very hard to check for duplications, etc.)
5 - Count mail-in votes without careful signature or registration verification.
   (This makes mail-in an easier choice for manipulators.)
6 - Allow votes to count that are received after Election day.
   (This can direct where mail-in votes are needed to go.)
7 - Stop vote counting for several hours before the final tabulations.
   (This allows for an assessment of how many votes are "needed" etc.)
8 - Do not allow genuine oversight of voting tabulation.
   (This would make it easier to lose or miscalculate actual votes.)
9 - Connect voting machines or precincts to the Internet.
   (This makes it quite easy for third parties to access and change votes.)
10-Distribute manipulations over multiple counties.
   (This makes the adjustments more difficult to find.)
11-Use multiple tactics to make manipulations.
   (This also makes the changes more difficult to find.)
12-Make most of the manipulations in unexpected districts.
   (In other words don't do as much manipulation where it's expected.)

There are undoubtedly more strategies those who are trying to control our politics would employ — but this is a representative sample. It should also be clear that many of these are difficult to find.

Frequently there is documented proof of some of these voting actions (e.g. leaving non-eligible voters on the rolls). However, these are usually dismissed with cursory responses such as: *we're doing the best that we can*, or *these deviations are not statistically significant*, or *our rolls are as accurate as other states*, or *there are some benefits for doing this (e.g. #3 & #6)*, etc.

However, studies like [this](#) and reports like [this](#) do not instill confidence that election results actually reflect the wishes of actual citizens.

So what can we do as scientists? Clearly we can't verify the legitimacy of every Pennsylvania vote submitted. On the other hand we can (from a scientific perspective and with sufficient data) provide a statistically strong assessment that reported votes in certain locations are statistically unusual. Such a determination should be treated as an indication that some type of accident or purposeful manipulation almost certainly occurred.

Such a science-based statistical analysis can not identify exactly what happened — or prove that fraud was involved. Honest mistakes, unintentional computer glitches, etc. can and do happen.

We approached this project assigning different experts to look at the Pennsylvania data from different perspectives. By-and-large the experts worked mostly independently of each other. As a result, there may be some overlaps in the analyses in the following five "chapters."

All of the experts agreed that there were major statistical aberrations in some of the Pennsylvania results, that are extremely unlikely to occur naturally.

Using more conventional statistical analyses, we identified eleven (11) counties with abnormal results (see Chapter 2). Due to time, data and manpower limitations, for this Report we focused on the statistical analysis for the worst five (5) counties. **Our strong recommendation is that each of those five Pennsylvania counties has an audited recount.**

If the results of a carefully audited recount are that there is **no** significant change in voting results for all of these five counties (very unlikely), then the authors of this Report recommend that we write off those county deviations as an extreme statical fluke, and that the Pennsylvania voting results be certified.

On the other hand, if the results of a carefully audited recount are that there **are** significant changes in voting results for some of these five counties, then the authors of this Report recommend that (as a minimum) that the next six (6) statistically suspicious counties also have an audited recount, prior to any certifying of the Pennsylvania voting results.

See ***Summary*** on the final page, for more conclusions.

— Editor, physicist John Droz, jr.

1 - Time Series Analysis of
Trump and Biden Votes in Pennsylvania

Dr. Louis Anthony Cox, jr.

As shown in Figure 1, data on cumulative counts for Trump and Biden in PA over the course of three days from November 4 to November 7 started with Trump ahead by more than 0.5M (by 540,522) at 11:00 AM on November 4 (time "0" on the left side of Figure 1). By 11:29 AM on November 7 (right end of Figure 1), the Biden curve had caught up with, and slightly exceeded (by 34,202) the Trump curve, with values at that time of 3,344,528 for Biden and 3,310,326 for Trump. The Biden count curve thus starts about 18% below the Trump count curve and ends up being about 1% above it (34202/3310326 = 0.0103). Even without detailed analysis, it is visually clear that the final values are remarkably close. This invites the question of whether such a coincidence indicates external intervention to close the initial gap between the curves, or whether it might plausibly have occurred without external intervention.
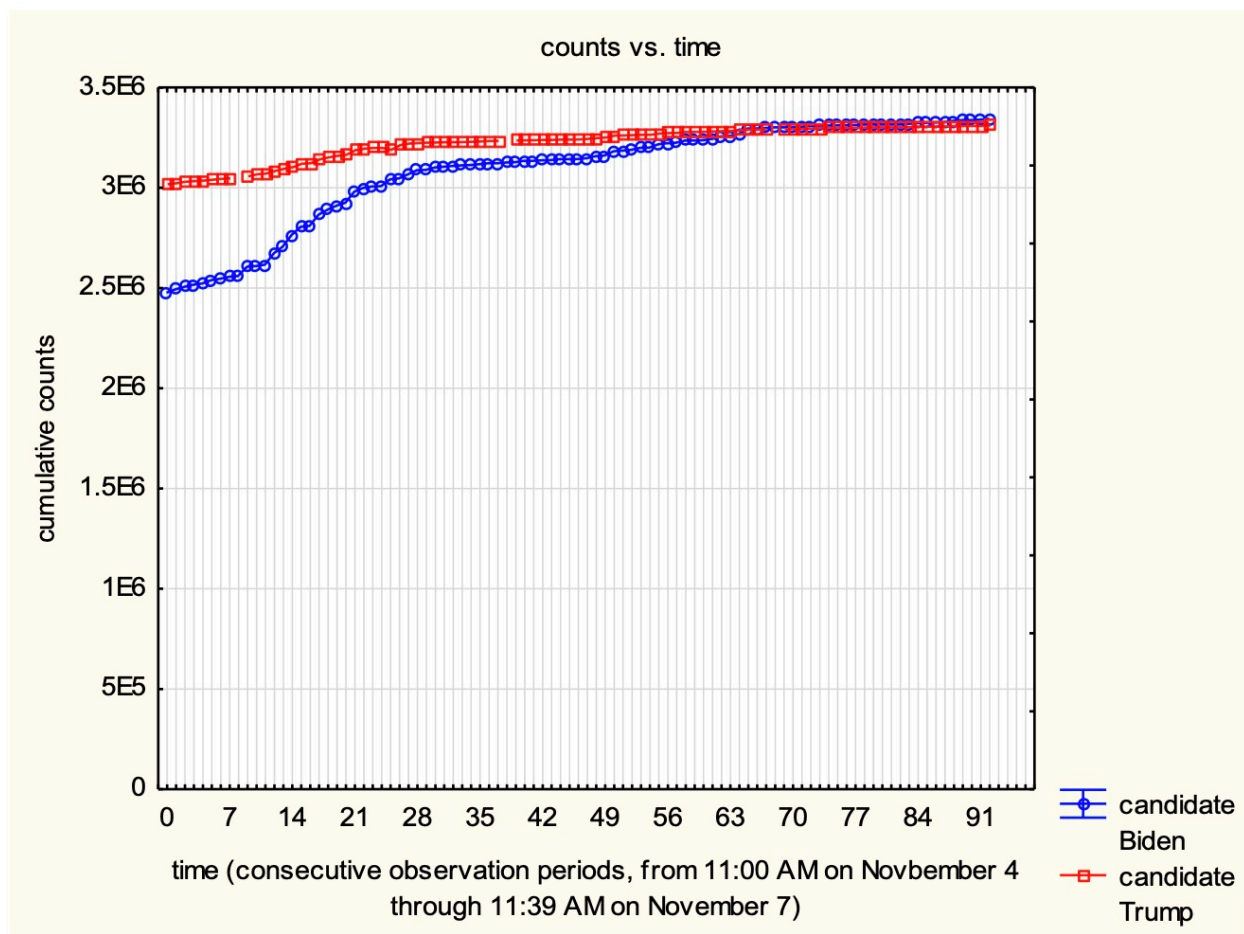

Figure 1. Time courses of Biden and Trump counts in Pennsylvania
from 11:00 AM November 3 *to* 11:29 AM November 7, 2020

How likely it is that such a near-coincidence of final counts (with the Biden curve finishing within about 1% of the Trump curve) would occur in the absence of external interference that brings the two curves together so closely? Although history never reveals its alternatives, computational statistics can help to determine what is plausible.   Figure 2 shows the approximate frequency distribution (histogram) of increments for Biden counts from period to period, with most being relatively small (left bar) but a few being an order of magnitude greater (right bar).
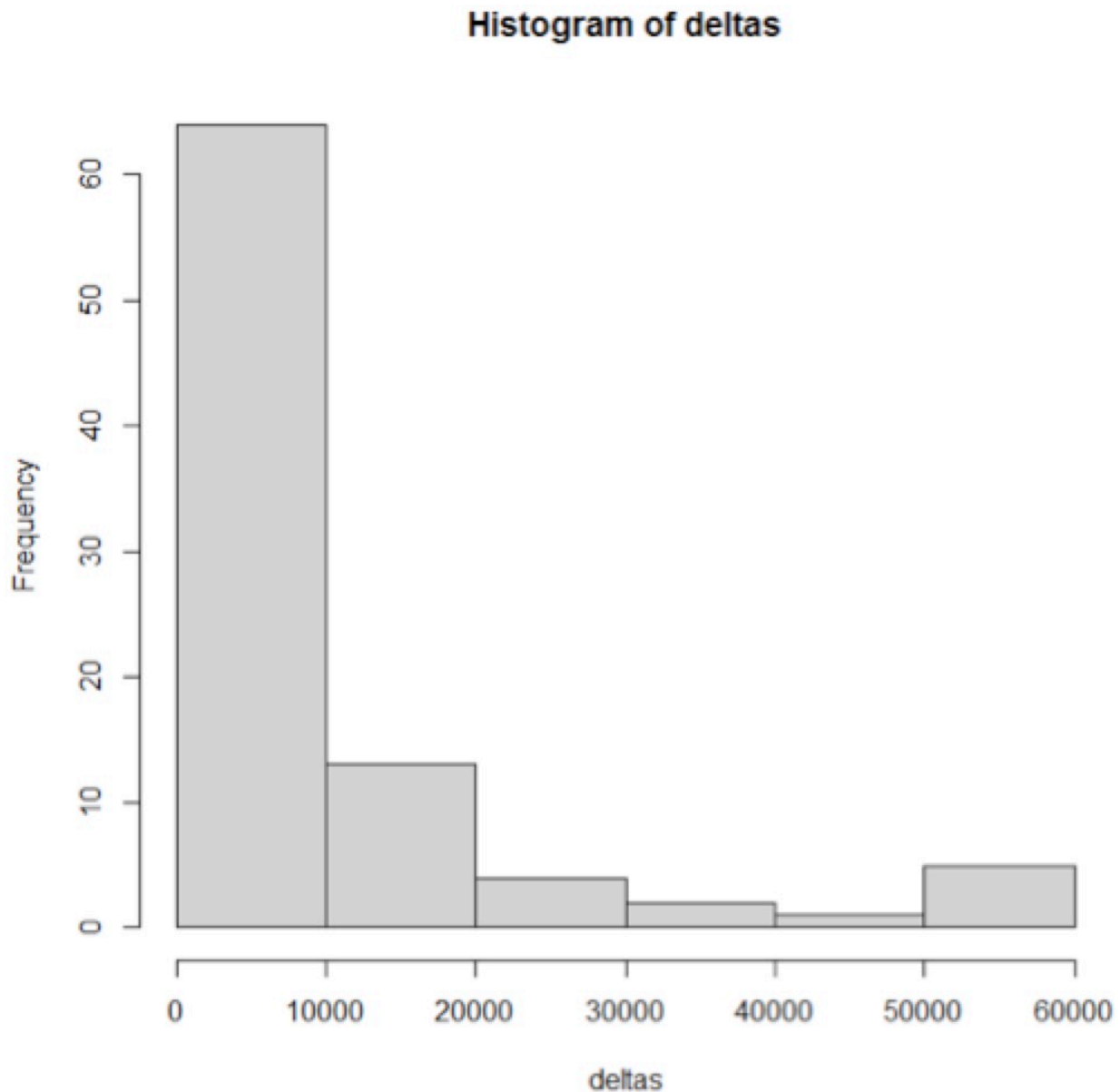

Figure 2. Histogram of deltas (increments between consecutive periods) of Biden counts

Randomly sampling from the distribution of increment sizes many times – a technique called "resampling" – and studying how much the sum of the increments varies across many random resampling scenarios provides one way to gain insight into whether the pattern seen in Figure 1 is unusual enough to indicate likely intervention. Figure 3 show the results of this statistical "bootstrapping" procedure for 10,000 randomly generated resampled ("bootstrapped") samples from the original data.

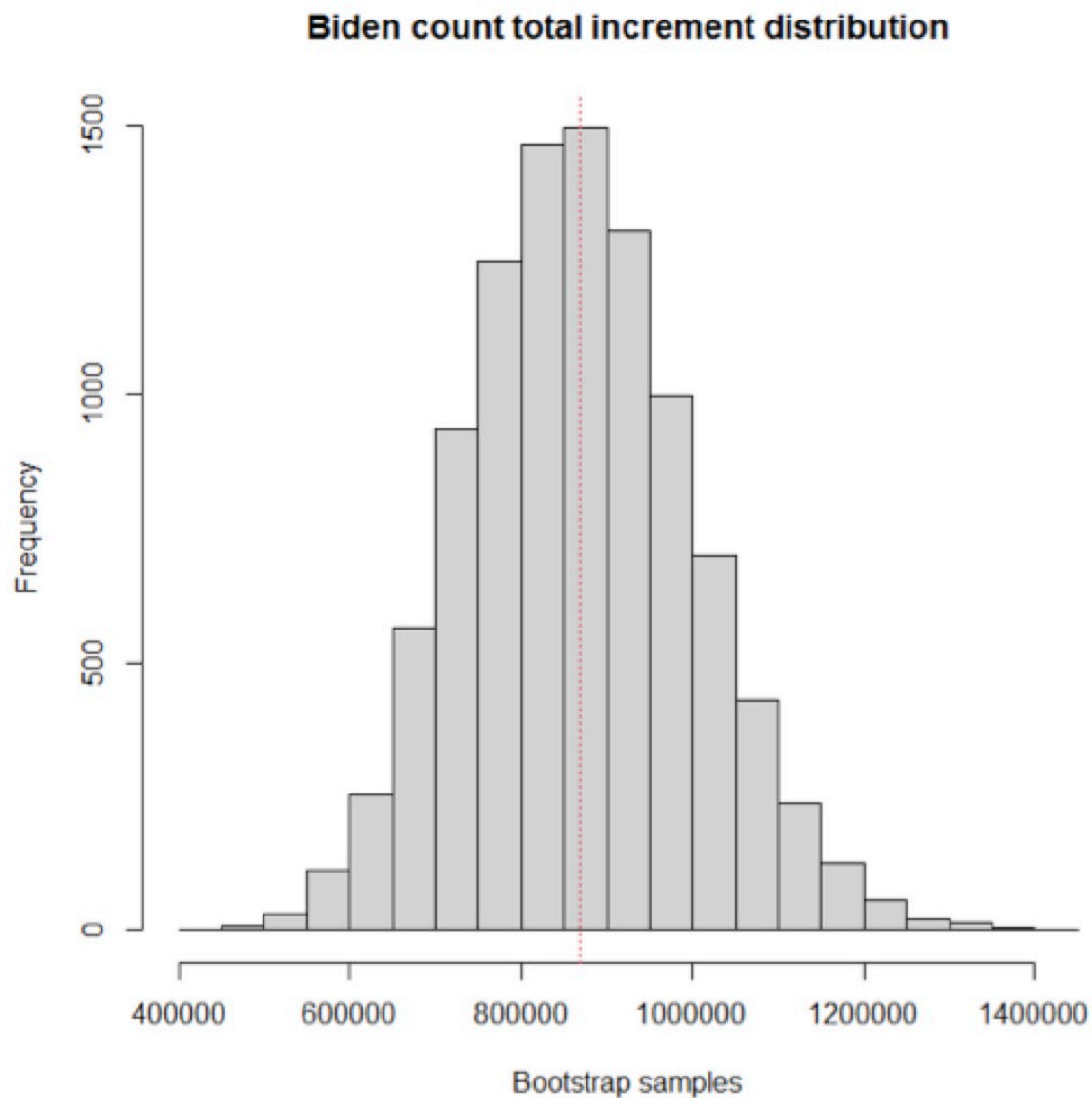## Biden count total increment distribution

Figure 3. Resampling (using the "bootstrap" method) shows that the sum of 90 increments sampled from the frequency distribution of increments observed in the Biden count time series (see Figure 2) spans a relatively wide range (roughly 3-fold). This makes it unlikely that the time course of Biden counts would end up within 1% of a specific value (here, the Trump final count) by chance.

Figure 3 shows that the total increment in Biden counts over the three-day observation period (modeled as the sum of about 90 consecutive increments) could plausibly have fallen anywhere in a fairly wide range, from less than 600,000 to more than 1,200,000, given the frequency distribution of increment sizes reflected in Figure 2. The probability of the final value falling within about 1% (34,202) of the final Trump value by chance alone is very small.

Conclusion: These calculations deliberately ignore the time patterns in the data (see Figure 1) to focus instead on the variability in the data. Based on this variability, it is not probable that the final Biden count would end up being extremely close (within about 1%) of the final Trump count by chance alone. The two final counts would be expected to differ by more if third parties had no mechanism for tracking or adjusting the Biden counts to the Trump counts.

2 - Pennsylvania County Voting Anomalies

S. Stanley Young, PhD, FASA, FAAAS

This report looks at Pennsylvania county voting, 2008 to 2020. The data set has 67 rows, with one row for each county. The first few rows are given here.

| RowID	PA Counties	Obama 2008	Obama 2012	Clinton 2016	Biden2020F
1	Adams	17633	15091	14219	17919
2	Allegheny	373153	352687	367617	415737
3	Armstrong	11138	9045	7178	8352

This report is in the form of text describing an item of interest with figures and tables along with discussion.

Summary:
- Philadelphia and Allegheny Counties are deviant in several respects including: they have high Democratic registration; they have a high percentage of voter turnout; the fraction voting changes dramatically from year to year; etc.
- The high vote for Biden counties are doubly unusual (i.e. are outliers) relative to previous presidential elections and relative to the remaining PA counties. Eleven such counties were identified. Together they report an excess of ~299,000 votes over expectation. The top five report about 216,000 votes over expectation. These increases in vote counts are statistically unusual, as most counties provide similar vote counts from Presidential election to Presidential election.
- Among the majority of PA counties, Biden's total was 70%± of registered Democratic voters. Among the ten anomaly counties (after elimination of Allegheny), Biden's total was 101%± of registered Democratic voters. That differential is highly suspicious.
- It makes sense to carefully evaluate the results for the 11 counties that have large increases in votes — i.e have an audited recount. Attention should focus on the top five problematic PA counties.

Item 1 —
Given in the figure on the next page are the change in voting for Biden 2020 relative to the average of three previous presidential elections (I'm calling that Dif1). The differences are ranked and plotted against the size of this difference. The largest increase is on the left and the largest decrease is on the right.

Bivariate Fit of Dif 1 By Rank1
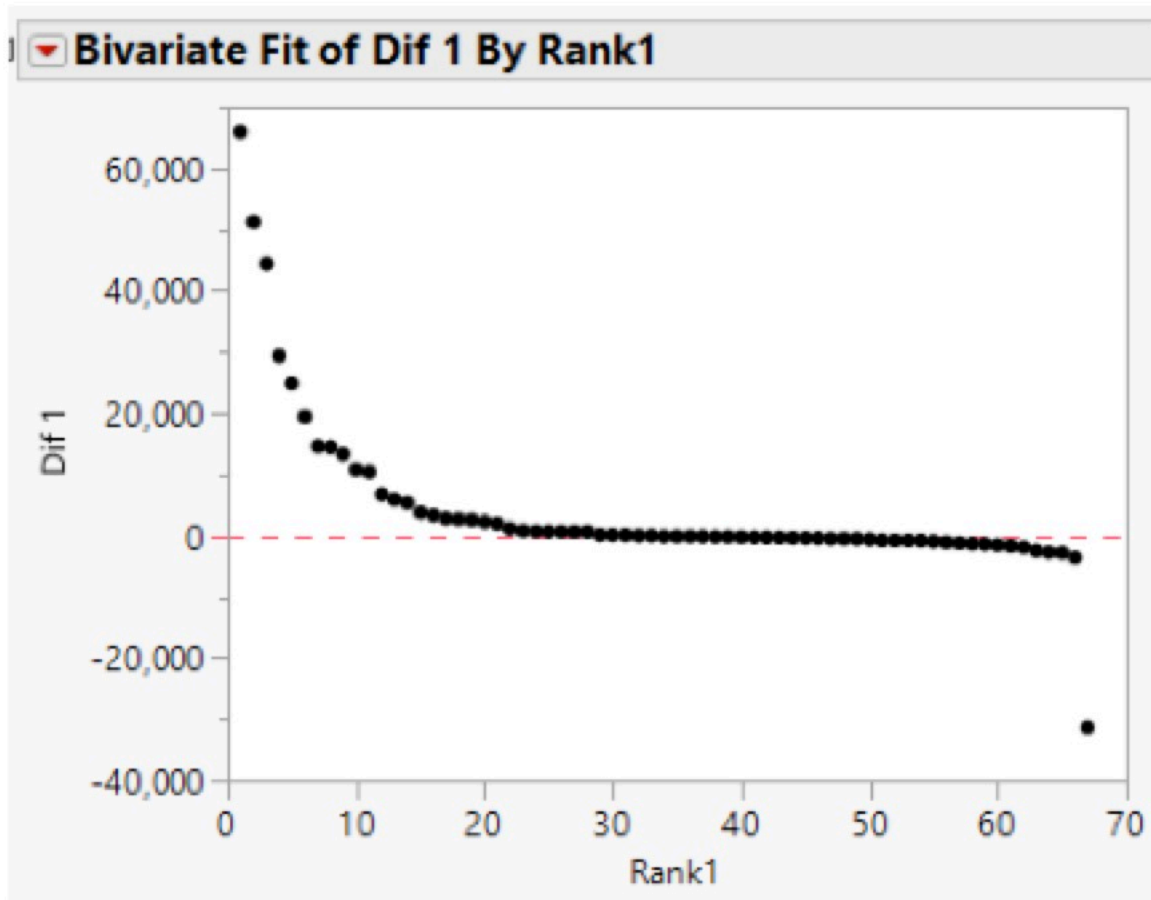
On the righthand side of the figure we see there are some counties where Biden did not do as well as the average. (The rightmost data point is Philadelphia which is a special case and will be covered elsewhere.) Toward the center of the figure we see that there was essentially no change from Biden to the average. It is common for people and counties to vote rather consistently from year to year. At the left side of the figure we see a slight rise, Rank 12 to Rank 22±, which is sort of a mirror image to the far right. The points from Rank 12 to Rank 66 are expected given the nature of voting – i.e. most people vote like they did last time.

The high values of vote counts, Ranks 1-11, on the left of the figure are substantially anomalous relative to the rest of the data. In the statistical literature they are called outliers – lying away from the body of the data. In these counties Biden did exceptionally well, while in majority of PA counties Biden did as expected (i.e. like previous elections). In some counties the Biden count is actually *lower* than previous Democratic presidential candidates. For 11 PA counties (the left most dots on the graph above, there are much larger increases in votes for Biden than are statistically expected.

Item 2 —
From the data in Item 1, the next page shows a list of the 11 outlier counties, where Montgomery County exhibits the most extreme statistical deviations.

PA Counties	Obama 2008	Obama 2012	Clinton 2016	Biden2020F	Dif 1	Rank1
Montgomery	253393	233356	256082	313543	65,932.7	1
Allegheny	373153	352687	367617	415737	51,251.3	2
Chester	137833	124311	141682	179065	44,456.3	3
Bucks	179031	160521	167060	198251	29,380.3	4
Delaware	178870	171792	177402	200911	24,889.7	5
Lancaster	99586	88481	91093	112536	19,482.7	6
Cumberland	48306	44367	47085	61168	14,582.0	7
Northampt...	75255	67606	66272	84145	14,434.0	8
Lehigh	87089	78283	81324	95539	13,307.0	9
Dauphin	69975	64965	64706	77387	10,838.3	10
York	82839	73191	68524	85323	10,471.7	11

As an example, consider Montgomery County. Obama / Hillary vote counts ranged from 233,000 to 256,000. Biden received 313,000. The eleven outlier counties together provide about 299,000 excess votes. The top five counties provide about 216,000 excess votes.
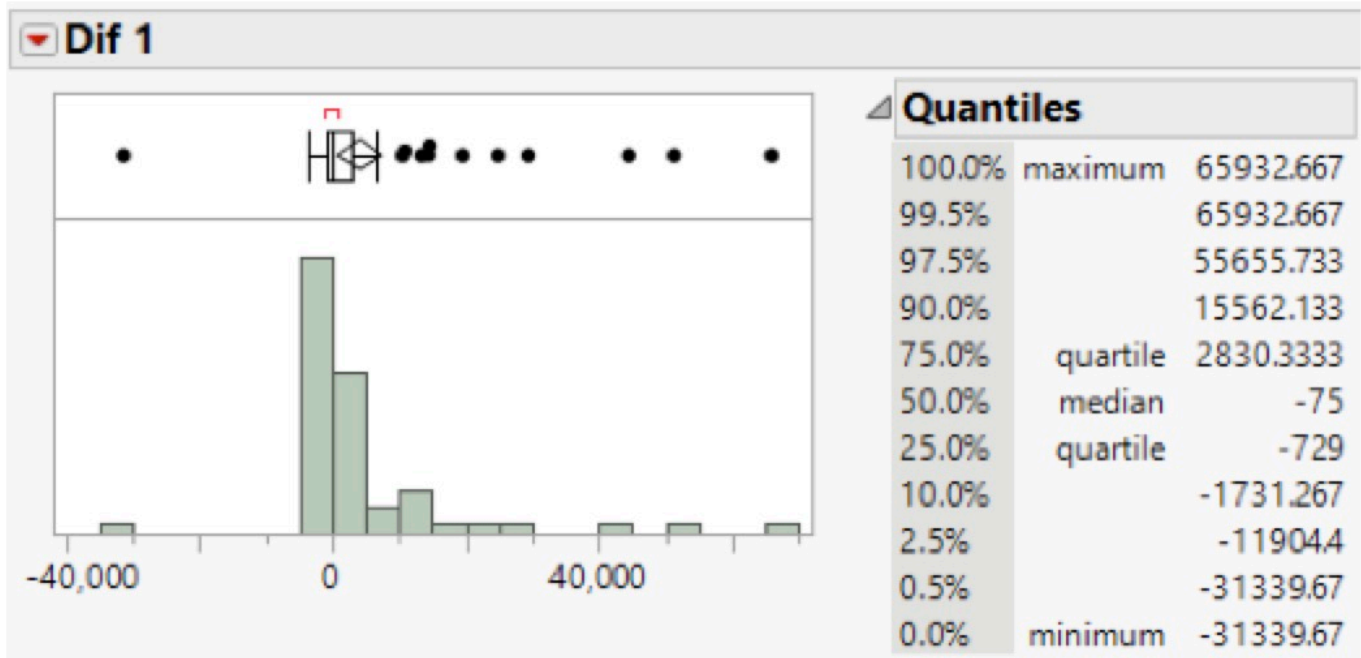
Item 3 —

The majority of PA counties (34) showed little change from previous presidential votes, i.e. little enthusiasm for Biden. We examine the bulk of the data, omitting for now those counties with a large increase, and Philadelphia in voting. We expect little change in the vote totals (DIF1) versus the average of previous votes and that is what we find for the bulk of the counties. In fact, there are more negative DIF1 values; note the large bar just below 0.
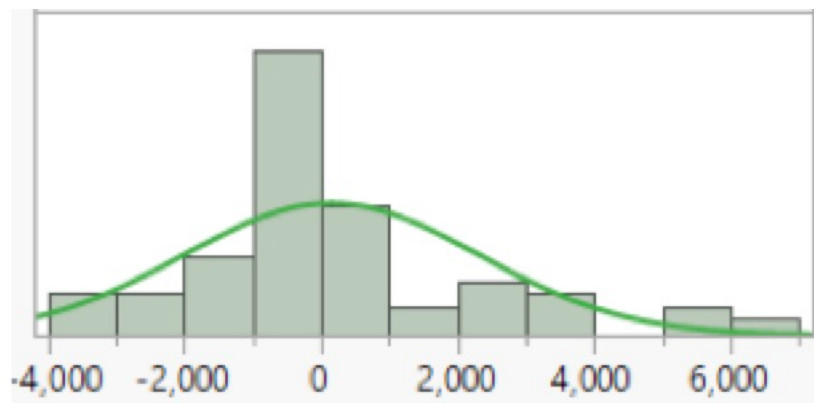
Item 4 —

We now look at the histogram for all the counties, including Philadelphia. (Philadelphia turned in 31,000± votes less than in the average of the prior three presidential elections.)


Dif 1

Quantiles		
100.0%	maximum	65932.667
99.5%		65932.667
97.5%		55655.733
90.0%		15562.133
75.0%	quartile	2830.3333
50.0%	median	-75
25.0%	quartile	-729
10.0%		-1731.267
2.5%		-11904.4
0.5%		-31339.67
0.0%	minimum	-31339.67

In the center of the figure, from -5,000 to about 6,000 we see bars that resemble a normal distribution; See Item 3. The values above 10,000 appear to be outliers. An outlier is an unusual number relative to other numbers in the collection. It is unusual to see a gain of 10,000 votes or more; reexamine Item 1.

Item 5 —

The changes in vote counts from Obama 2008 to Obama 2012 were mostly negative, give here as Obama Dif and is plotted against their ranks (next page). The votes for Obama were high in 2008. Most counties provided fewer votes in 2012, the down sloping set of points. At the end of this down-sloping drift, there are dramatic falls in vote counts, outliers of votes lost.

RowID	PA Counties	Obama 2008	Obama 2012	Obama Dif	Rank O Dif
6	Berks	97047	83011	-14036	64
9	Bucks	179031	160521	-18510	65
46	Montgomery	253393	233356	-20037	66
2	Allegheny	373153	352687	-20466	67

Bivariate Fit of Obama Dif By Rank O Dif
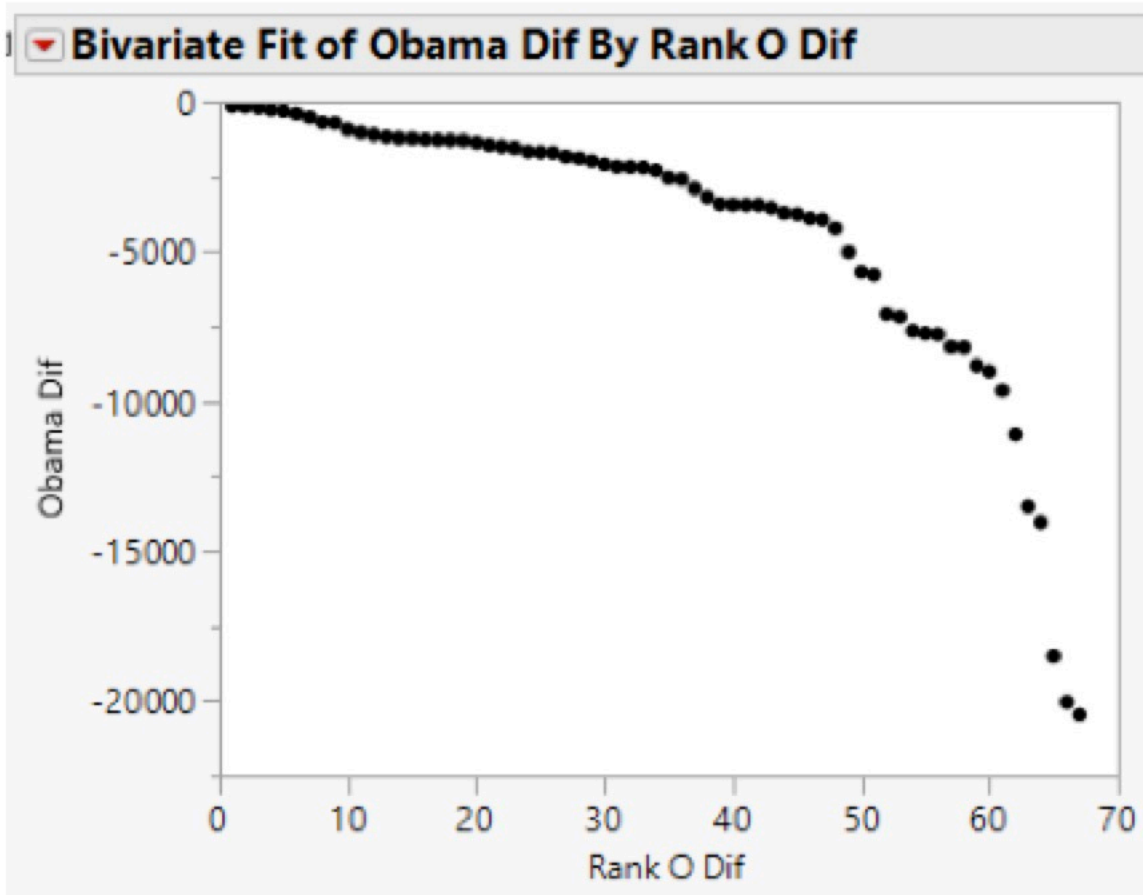
It is curious that many of the same counties, e.g. Montgomery and Allegheny, come up having large declines with Obama 2012, but having large increases with Biden 2020. These wild swings are extremely unusual as most counties, where voters vote similarly over time.
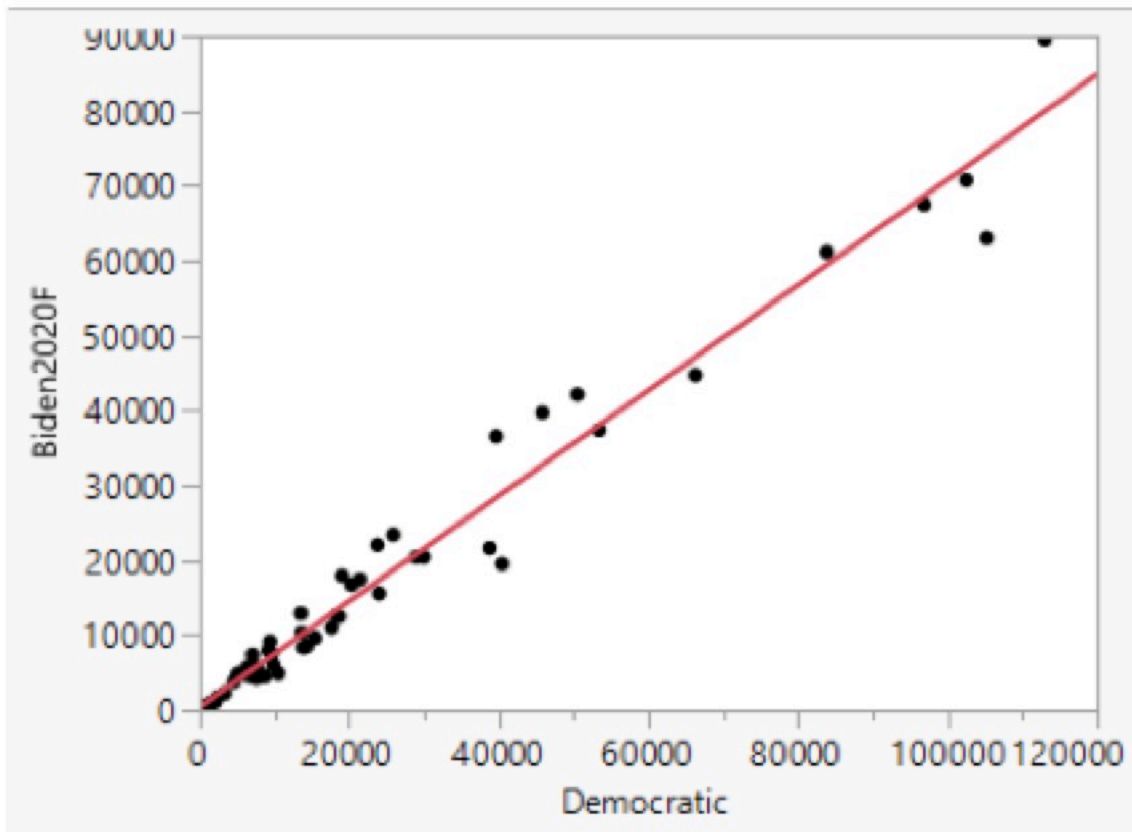
Item 6 —
We seek to estimate the fraction of registered Democratic voters that voted. We want an unbiased estimate, so the 11 outlier counties and Philadelphia were removed from the analysis. 55 PA counties were used for simple linear regression.
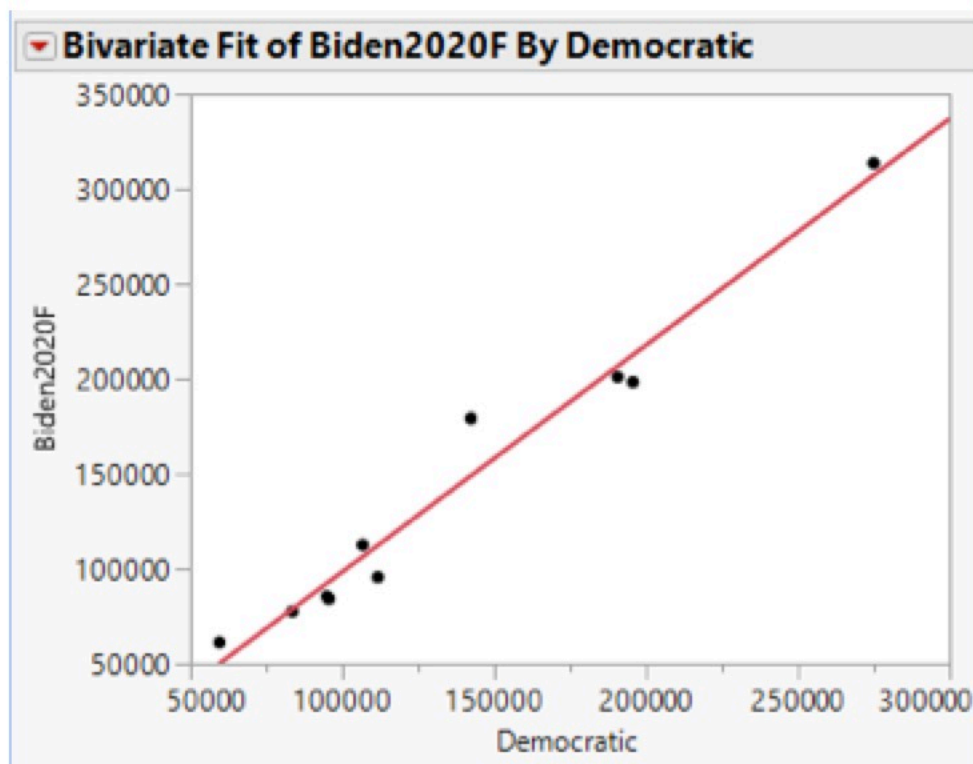
The data are fit well with a simple line (see next page)

      Biden2020 = 439.8738 + 0.7036542*Democratic

This means that we expect 70%± of registered Democratic voters to vote in normal (the majority of) Pennsylvania counties.

Item 7 —

We seek to estimate the fraction of registered Democratic voters that voted among the outlier counties. We want an unbiased estimate, so we removed Allegheny and Philadelphia counties as they are rather unique. Ten counties were used for simple linear regression.

The data are fit well with a simple line (see prior page).

$$Biden2020 = -21215.45 + 1.1943149*Democratic$$

This means that the number of Biden votes in ten of the outlier counties was **101%±** of registered Democratic voters (<u>vs</u> the majority of other PA counties where it was **70%±** — an extraordinary statistical difference). That is not logical or reasonably explainable legally. The most likely explanation is that excess votes were added to the Biden total that did not come from voters.

Item 8 —
Our goal here is to estimate the expected relationship of Biden votes to the number of registered Democrats. There are non-problematic counties (55) and there are problematic counties (11 – 1 = 10). Note that Philadelphia and Allegheny counties are omitted. We also want to know the number of actual Biden votes per registered Democrat, separately for non-problematic and problematic counties. We use two methods of simple linear regression. More standard is the *Intercept Model* linear regression. In this method a line is placed through the data without constraint, the line can move and twist. Less standard is the *No Intercept* method. In this method the line is constrained to go through zero on the Y and X axes. Either method can make sense, so we present both. We focus on the slope of each of the four models, Intercept/No intercept, Non-problematic/Problematic. The slope indicates the number of Biden votes expected per registered Democrat voter. Here are the four slopes.

N	55	10
	Non-Problematic	Problematic
Intercept Model	0.7037	1.1943
No Intercept Model	0.7114	1.0654

First consider the 55 non-problematic counties. These are the counties where we did not find evidence of voting problems. The slopes for the two models are quite similar and indicate that **for every 100 increase of registered Democrat voters, there should be a 70± vote increase for Democrats.**

Both slopes for problematic counties are much larger and rather different from each other. That both are over 1.0 indicates that for every 100 registered Democrat voters there are more than 100 Democrat votes, which is quite improbable. The Intercept Model is **not** constrained to pass through 0,0 so it has more freedom to fit the data. Its slope is greater and indicates 119 Democrat votes are occurring for each 100 registered Democrats, again improbable.

The No Intercept Model is constrained to pass through the 0,0 point. With either model, the problematic counties give an improbable result, more Biden votes than there are *registered* (not voting) Democrat voters.

Next, we compute the actual number of Biden votes per registered voter.

	Non-Problematic	Problematic
Actual votes per Reg Dem	72.8531	101.0012

We see that in non-problematic counties that an average of about 72 votes are obtained for each 100 registered voters, which comports with usual voter history. For problematic counties we get an average of 101 voters per 100 voters, which is quite unusual. It is instructive to see the actual data.

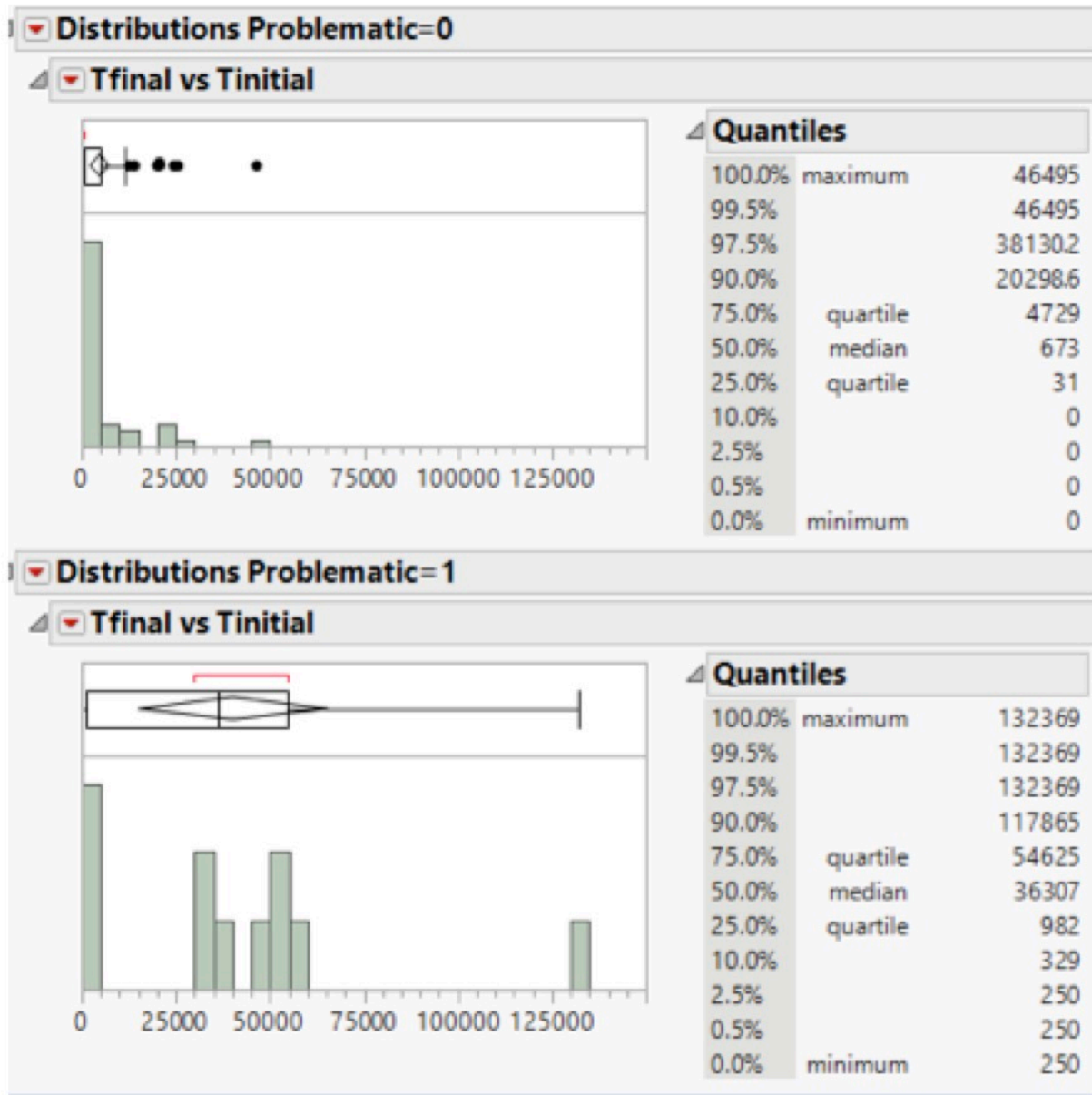| RowID	PA Counties	Biden2020F	Democratic	%Democrat
46	Montgomery	313543	274955	114.0
2	Allegheny	415737	555649	74.8
15	Chester	179065	142423	125.7
9	Bucks	198251	195772	101.3
23	Delaware	200911	190702	105.4
36	Lancaster	112536	106762	105.4
21	Cumberland	61168	59656	102.5
48	Northampt...	84145	95710	87.9
39	Lehigh	95539	111803	85.5
22	Dauphin	77387	83635	92.5
67	York	85323	95027	89.8

Item 9 —
Vote counts were secured for Wed, Nov 4 and also the final counts. The difference between these counts is the number of mail-in votes. Here we examine the distribution of those votes between problematic and non-problematic PA counties. **Many** more votes were added to the problematic counties compared to the non-problematic counties. As we often do, Philadelphia was not included in the following analysis.

Non-problematic counties (0) added a median of <u>673</u> votes per county.
Problematic counties (1) added a median of <u>36,307</u> votes per county.

The number of mail-in votes in non-problematic counties can serve as a proxy for "voting/business as usual". The mail-in vote for problematic counties can be taken as another aspect of the problematic nature of these counties

Distributions Problematic=0

Tfinal vs Tinitial

Quantiles

100.0%	maximum	46495
99.5%		46495
97.5%		38130.2
90.0%		20298.6
75.0%	quartile	4729
50.0%	median	673
25.0%	quartile	31
10.0%		0
2.5%		0
0.5%		0
0.0%	minimum	0

Distributions Problematic=1

Tfinal vs Tinitial

Quantiles

100.0%	maximum	132369
99.5%		132369
97.5%		132369
90.0%		117865
75.0%	quartile	54625
50.0%	median	36307
25.0%	quartile	982
10.0%		329
2.5%		250
0.5%		250
0.0%	minimum	250

3 - A Testable Hypothesis of Fraud using a Predictive Model in the Pennsylvania 2020 Presidential Vote for Montgomery County

(Condensed Version)

Dr. Samuel Culper III, Dr. Nathan Hale, Dr. Abraham Woodhull

Contents

Executive Summary

Analysis – Statistical analysis from a coordinated team of (masked for this report) experts in the field applied classical statistical methods show irregularities in not one, but many major counties in the Pennsylvania 2020 Presidential Election. The analysis couples with extremely large turnouts in excess of 80%, which significantly exceeds similar democrat cities which average around 64%, Philadelphia included. This turnout is also in excess of the same counties' turnout in the 2008 Obama election. These facts suggest a mathematically extraordinary event occurring in multiple counties simultaneously at a magnitude well above what is needed to change which candidate won Pennsylvania's electoral votes.

Hypothesis – A predictive model hypothesizes up to 27,000+ votes fraudulent in Montgomery County alone, with the model able to expand to other counties with more time. Biden's votes alone are being inflated. The hypothesis suggests a singular or small set of actors in a position to intercept and modify all precincts data applied a "stuffing the tail" vote fraud scheme to increase Biden's votes in democrat heavy districts that would be undetected by the workers at the precincts but exceedingly trivial to detect mathematically.

Testable Proof – Search for forensic voter irregularities in a small selection of unrelated districts in Montgomery County from the list of specific predictions. Not all counties are predicted to have been defrauded. After an audit of the district, the hypothesis should be tested by calculating a quantitative sum of voting irregularities and compared against the prediction value. If the quantitative comparison matches or is reasonably close to the prediction of that district, and proven additionally true across the other unrelated precincts, then the hypothesis is proven true with statistical confidence that voter fraud has occurred at scale.

A more detailed version of this report on the analysis, explanation, and prediction model are available as needed.

Mathematical Evidence of Fraudulent Activity

The Biden vote change over Hillary Clinton's vote in 2016 in Montgomery County, PA shows a mathematically fraudulent result that is adding votes unnaturally to the tail of the distribution. High turnout or massive changes in voter preferences cannot statistically create this result in any reasonable probability. The signature is mathematical evidence of a classical form of fraud called "stuffing the curve". The 2008 sub-prime mortgage risk management meltdown is a recent example of this kind of fraud.

Montgomery Biden Vote Gain vs 2016 per Precinct
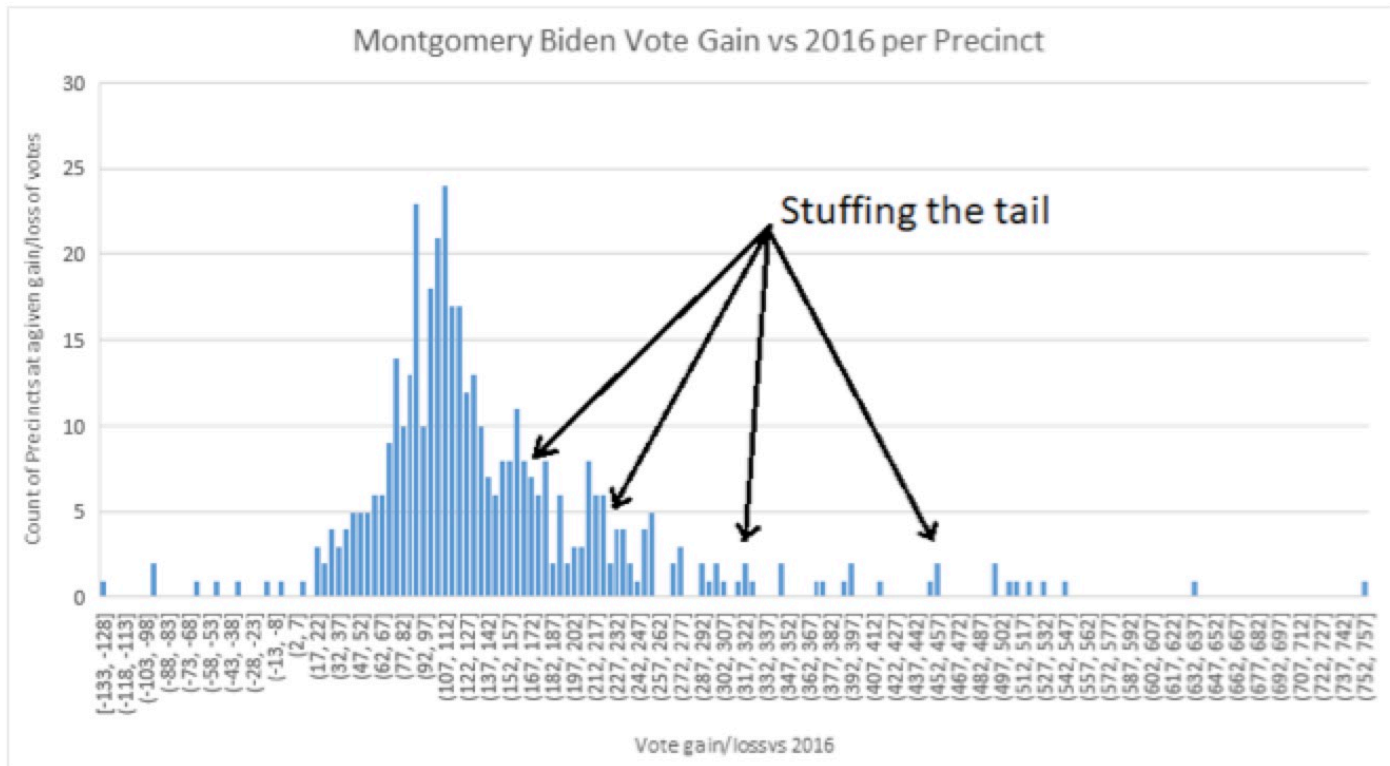
Specifically, an actor is adding Biden votes in heavy Democratic districts in excess of the real result with the assumption the addition will not be noticed due to the already high Dem/Rep ratio.

The "stuffing the tail" effect is seen in the following average vote gain over 2016 per precinct.

	Curve MEAN	
Trump		50.45
Biden		143.30
Diff		92.85
2020 Dem/Rep Ratio		2.84
%	74D/26R	
2016 D/R Ratio		1.57
%	58D/37R	

This gain in votes for both Democrat and Republicans is not only representative of increased turnout without any losses to their 2016 take, but an astonishing 16-point increase in the take of new votes favoring Biden, requiring not only a matched turnout of supporters to keep pace with Trump turnout at a higher required multiple, but anomalous amounts of additional new turnout on top. In a large set of precincts the gains by Biden alone exceed 100% of the precinct's new registrations.

A simple mathematical model hypothesizing voter fraud at a en masse level is constructed to remove the presumed fraudulent additions and identifies down to the precinct level the expected amount of voter irregularities to be found. The hypothesis may be proven or disproven by multiple tests at unrelated precincts and comparing against forensic evidence.

Start here – Lower Merion 12-3

The following screenshots are taken 2:40 AM 11/14/2020 on the Montgomery County results website. The model currently predicts the following amount of fraud in this district:

		Fraud	
Actual datapoint			
Predicted Total			
Municipal	**District**	**Count**	
Lower Merion	12,3		11

The next screenshot shows this sub district violating by at least 15 votes with the gross, simple prediction provided.
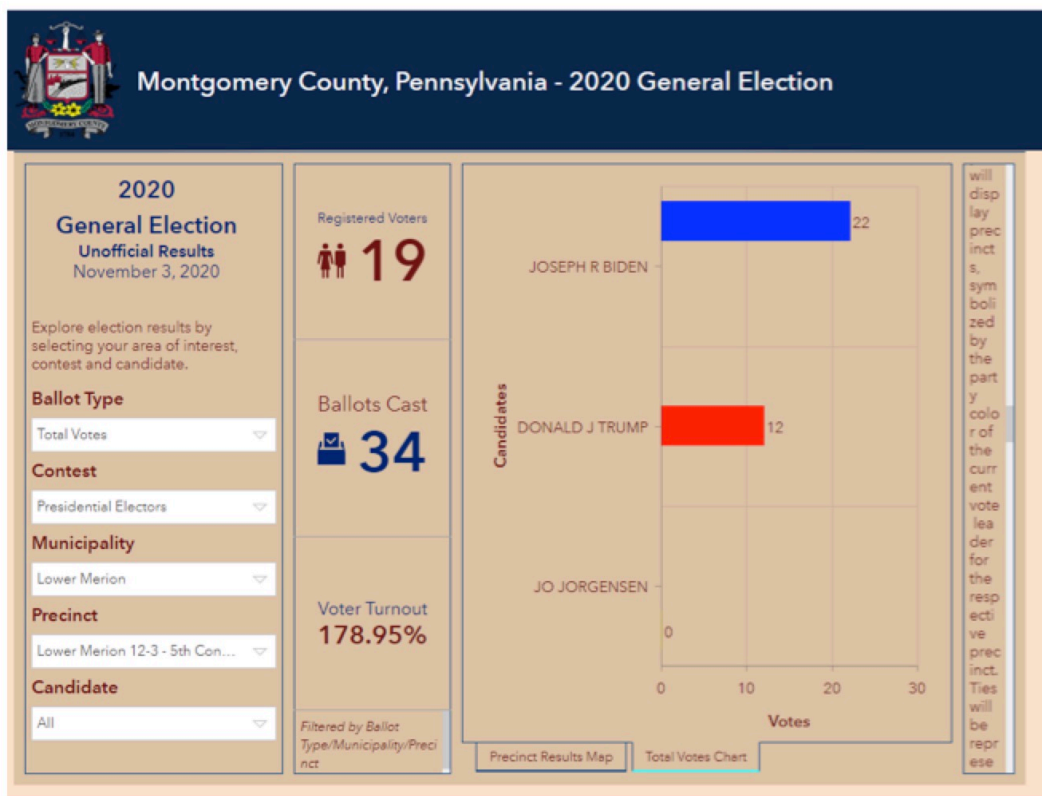
Therefore, <u>one point has already been verified</u>, and this precinct should be examined to explain the obvious "database error" (a.k.a. "fraud" that was too small to be noticed).

The Montgomery Fraud Prediction List Per Precinct

The predicted fraudulent districts at a ratio according to 2016 Dem/Rep distributions, using Trump's vote as the constant for turnout. Only the top few are listed here, and the full list is available on request. The net predicted set of fraudulent votes that may be tested in Montgomery County is roughly 27,000 votes for Biden alone.

Sorted Prediction

Municipal	District	Predicted Fraud Count
Trappe Boro	1	353
Upper Providence	Oaks	325
Upper Pottsgrove	1	309
West Norriton	3,1	305
Upper Providence	M-C	281
Upper Providence	Trappe Boro	271
Perkiomen	1,6	268
Skippack	3	267
Lower Providence	3,3	266
Upper Providence	Mingo 2	263
Perkiomen	2	262
Skippack	2	262
Limerick	1	259

4 - A Testable Hypothesis of Fraud using a Predictive Model in the Pennsylvania 2020 Presidential Vote for Allegheny County

(*Condensed Version***)**

Dr. Samuel Culper III, Dr. Nathan Hale, Dr. Abraham Woodhull

Contents

Executive Summary

Analysis – Statistical analysis from a coordinated team of (masked for this report) experts in the field applied classical statistical methods show irregularities in not one, but many major counties in the Pennsylvania 2020 Presidential Election. The analysis couples with extremely large turnouts in excess of 80%, which significantly exceeds similar democrat cities which average around 64%, Philadelphia included. This turnout is also in excess of the same counties' turnout in the 2008 Obama election. These facts suggest a mathematically extraordinary event occurring in multiple counties simultaneously at a magnitude well above what is needed to change which candidate won Pennsylvania's electoral votes.

Hypothesis – A predictive model hypothesizes up to 30,500+ fraudulent votes in Allegheny County alone, with the model able to expand to other counties with more time. Biden's votes alone are being inflated. The hypothesis suggests a singular or small set of actors in a position to intercept and modify all precincts data applied a "stuffing the gaussian" (different variant on what happened in Montgomery – indicates possibility of a different actor) vote fraud scheme to increase Biden's votes in Democrat heavy districts that would be undetected by the workers at the precincts but exceedingly trivial to detect in a Gaussian differential.

Testable Proof – Search for forensic voter irregularities in a small selection of unrelated districts in Montgomery County from the list of specific predictions. Not all counties are predicted to have been defrauded. After an audit of the district, the hypothesis should be tested by calculating a quantitative sum of voting irregularities and compared against the prediction value. If the quantitative comparison is reasonably close to the prediction of that district, and proven additionally true across the other unrelated precincts, then the hypothesis is proven true with statistical confidence that voter fraud has occurred at scale.

For details on methods and descriptions, see the full version of the Montgomery County, PA vote prediction report. This report simply details the fraud specific to Allegheny County and where to look.

Mathematical Evidence of Fraudulent Activity

Allegheny County shows a slightly different fraudulent signature than Montgomery County. To see it clearly, we had to take out the outliers and see the proper shape of the distributions:
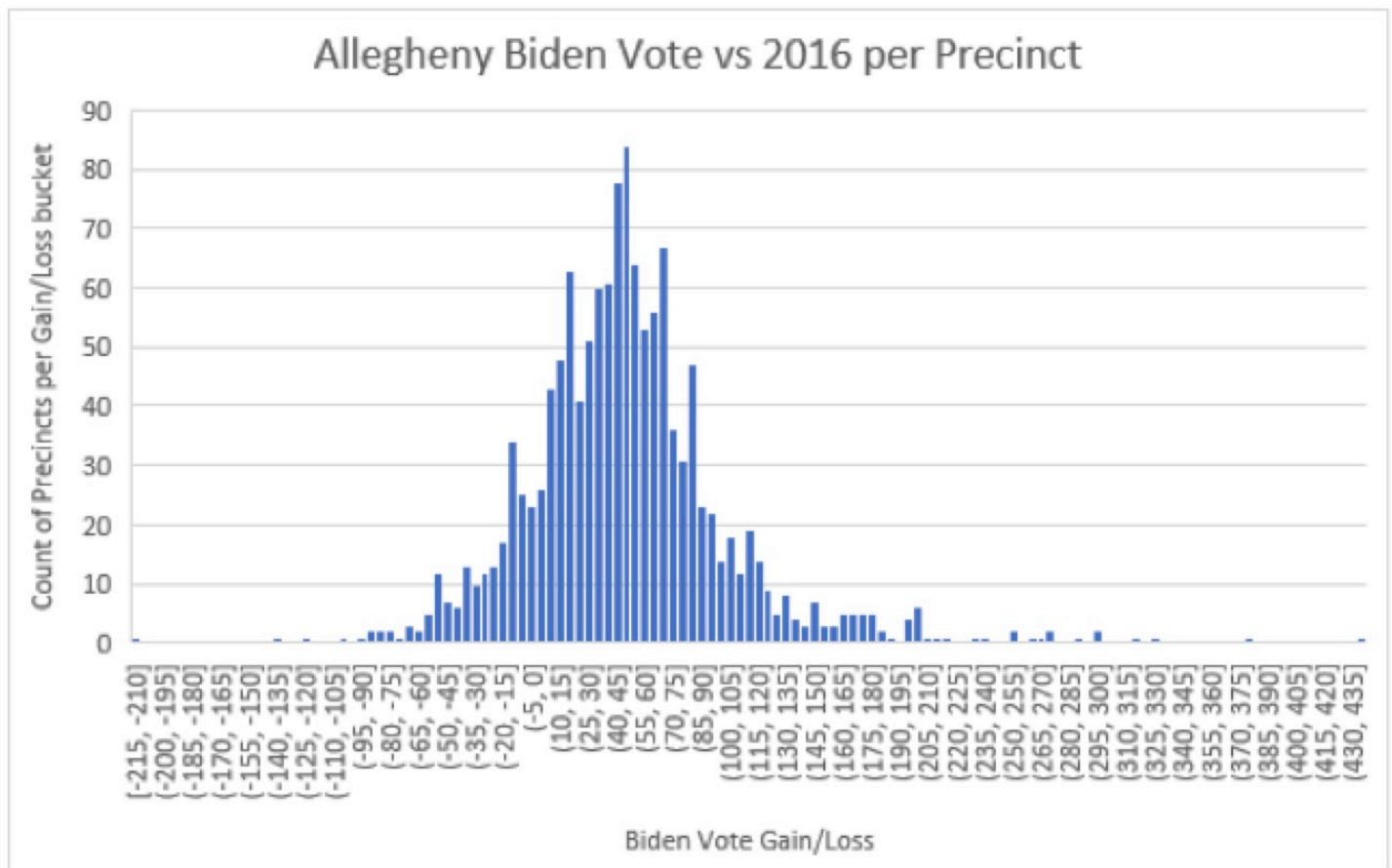
Initial analysis was a bit of a surprise, as instead of a "stuffing the tail" attack (as in Montgomery), this was a "stuffing the …. everything…?" kind of behavior. With all precincts game for adding Biden counts, the visualization above suggests it is not a straight multiplier (like Milwaukee: see Milwaukee vote fraud prediction report), and it must be additive.

After filtering out the bad tails and sorting by predicted precinct to again, a target of 2016 Dem/Rep ratios with Trump votes as the control for turnout matching, the answer for the temptation to "stuff the gaussian" was to cover for significant LOSSES in each precinct. In particular the greatest seeming outliers of the fraud prediction are those with almost 50/50 splits Dem/Rep, or slight advantage Republican.

Additionally, Pittsburgh seems to have resulted in some losses for Biden, so selective other districts seemed to plug the hole so as to appear not LOSE to votes locally from 2016.

Ratio target was as such for public data and correction ratio for the prediction. Additionally, as in the other predictors, the extreme MEAN difference of each curve above and beyond the 2016 total gave us a rough starting point on what to correct.

The county in 2016 had roughly a 59/41 (D/R) ratio, but the 2020 election was adding votes above the 2016 total at a ratio of 73/27 (D/R), indicative of the fraud and easy to identify at the precinct level base on each precinct's voter distribution history.

Total prediction hypothesizes that the 2020 election results, which look like this

2020	Register	Voted	Biden	Trump	D/R
	942851	722145	428876	282170	1.52
turnout	76.59%	share	59.39%	39.07%	

Are instead supposed to look like this:

Total Predicted 2020	Register	Voted	Biden	Trump	D/R	Fraud
	942851	681522	399341	282170	1.42	**30518**
turnout	72.28%	share	58.60%	41.40%		

which seems to still have a few uncorrected stuffings, but is much closer to the voter behavior differentials as seen in the Trump gaussian control.

The Allegheny Fraud Prediction List Per Precinct

Top cases listed. For full list, see the full Allegheny Fraud Prediction report. 692 / 1321 Precincts are predicted to have fraudulent votes:

Predicted Total

Precinct	Likely Fraud	D/R
MCCANDLESS WARD 6 DIST 3	263	0.97
PINE DIST 5	218	0.72
ROSS WARD 4 DIST 2	207	0.77
OHIO DIST 3	202	1.03
PINE DIST 2	188	0.59
OHIO DIST 2	177	0.75
FRANKLIN PK WARD 2 DIST 3	176	0.94
BETHEL PARK WARD 3 DIST 2	175	0.70
ROBINSON DIST 9	163	0.82
BETHEL PARK WARD 5 DIST 3	163	0.95
N FAYETTE DIST 2	155	0.81
PINE DIST 4	152	0.69

5 - Potential Voter Fraud in Pennsylvania

Dr. William M. Briggs

I used data provided to me of the hour-by-hour vote totals for both Biden and Trump beginning the day after the election. All analyses were conducted in R (version 3.6.1).

The following plots the cumulative total for both candidates beginning after election night.
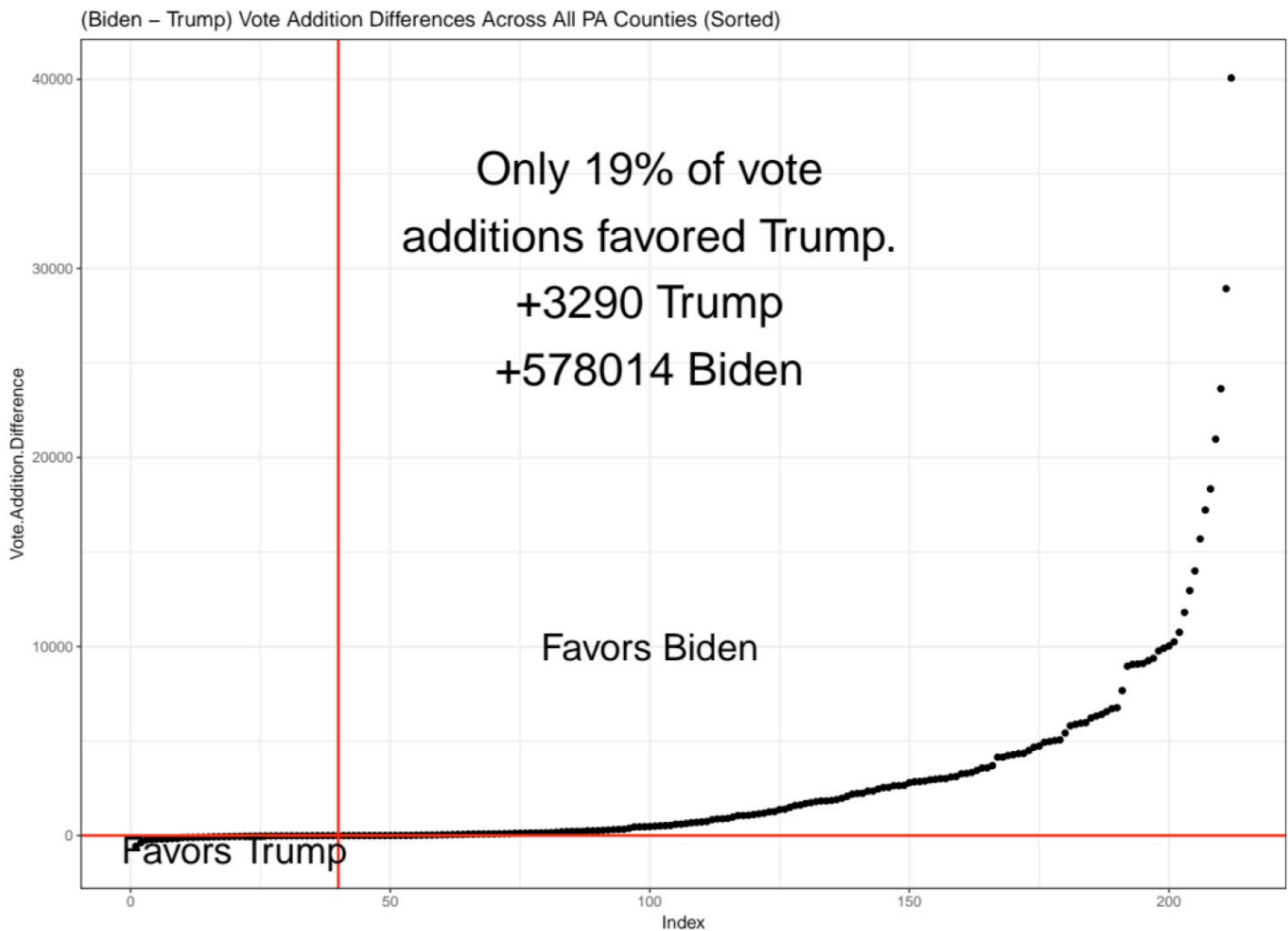
VOTE TOTALS

Running Candidate Totals of Late Votes Pennsylvania
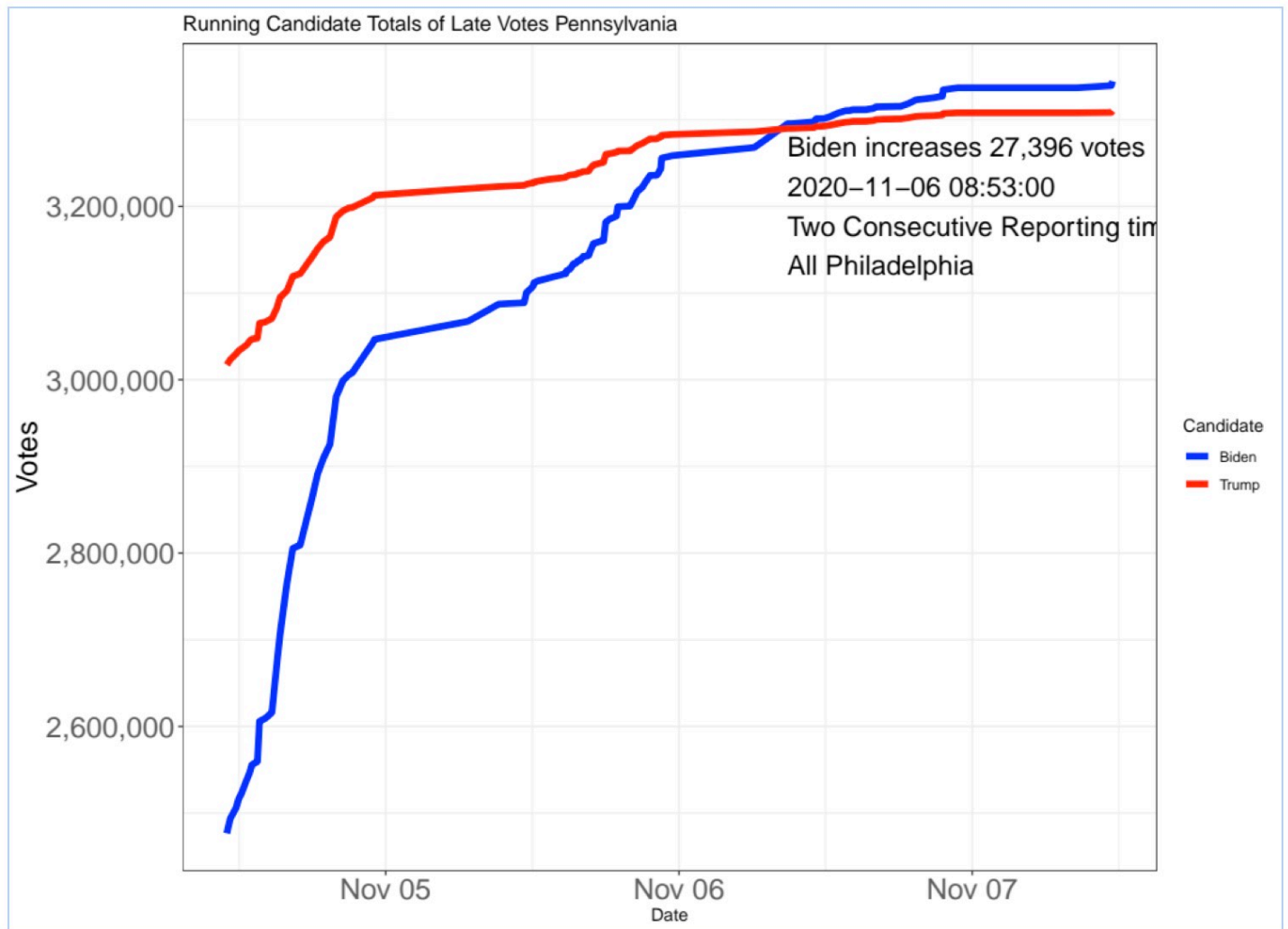
Trump starts well ahead, but due to enormous increases at specific time points (demonstrated next), Biden catches up rapidly. Obviously, those adding the votes in time do not know what the eventual total will be. This is what makes the late addition on the 6th suspicious. Biden's total was augmented by just over 27 thousand votes, which was just enough to put him ahead. The time was also near where the vote count was nearing its end.

(Biden – Trump) Vote Addition Differences By County and Time
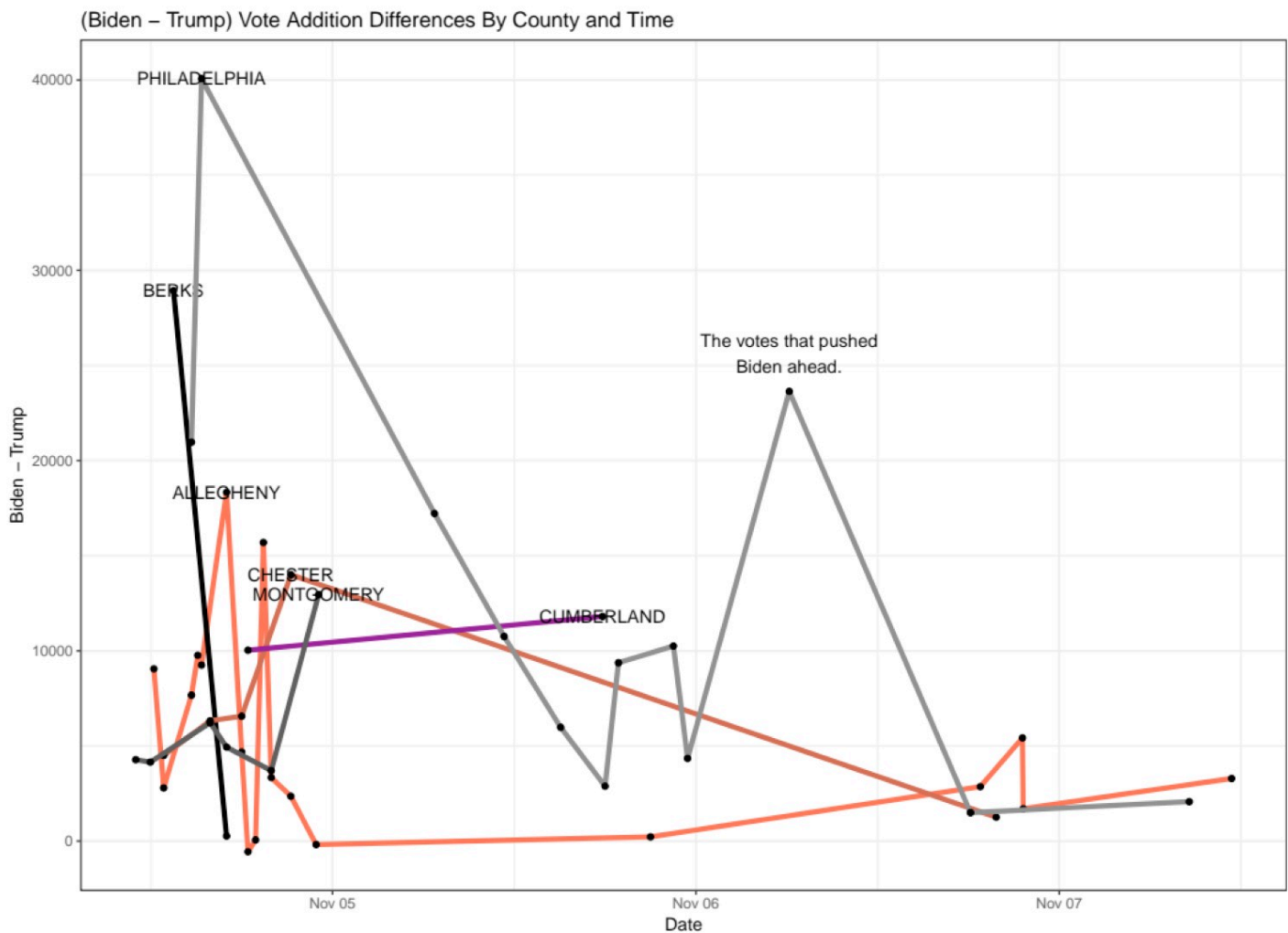
Pictured here are (Biden – Trump) vote differences in time for several counties (all with major additions to the counts). County names appear at the maximum of the difference. Berks, Philadelphia, Chester, Montgomery, Cumberland and Allegheny counties all give early advantage to Biden. But it was Philadelphia county that pushed Biden ahead. No other vote additions after this time were important or came close to changing the lead for Biden. The size of the difference at the late time bears investigation.

Most of the vote changes after election night favored Biden, which his odd. Here is a picture of these sorted from low to high (Biden – Trump) non-zero vote changes.

(Biden – Trump) Vote Addition Differences Across All PA Counties (Sorted)

Only 19% of vote additions favored Trump.
+3290 Trump
+578014 Biden

Favors Biden

Favors Trump

Only 19% of the times when new votes were tallied favored Trump, and for only an advantage of 3,290 votes. 81% of the changes favored Biden, for an advantage of over 550,000 votes. There is also a visible difference in distribution of these additions, centering (as the picture above shows) mainly on Philadelphia county.

This next plot (next page) makes this more apparent. It shows all additions for both candidates, sorted from the counties which added the most votes to the least. Blue dots are votes for Biden, red for Trump. Several counties are highlighted that show curious large additions for Biden.

Vote Additions > 0

CURIOUS COUNTIES

Here is a plot (using data on final election tallies provided by the same source) of the proportion of total votes Democrat presidential candidates received since Obama's first run. Those counties in which Biden improved over Obama's first run are highlighted in blue.

Proportion of Democrat Vote to Total Votes: Blue Biden Increase Over Obama 1
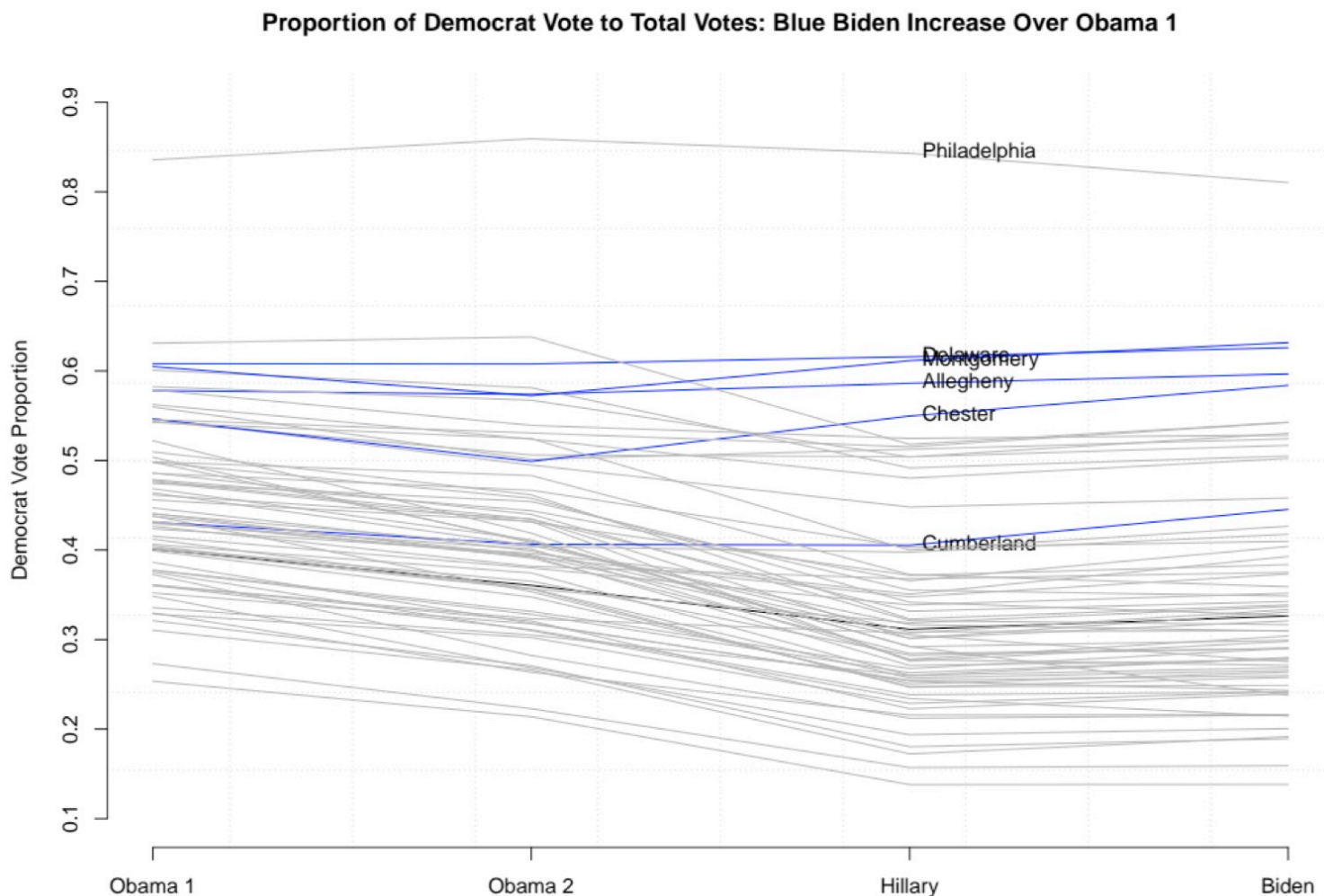
The proportion Democrats had been getting was declining steadily until 2020. Most stayed about the same from Hillary to Biden, but a few rose about their 2008 levels, which is odd, given Obama's gargantuan popular support at the time, and Biden's almost invisible public support in 2020.

The next picture is the same, but for total votes received for Democrat candidates.

Democrat Votes by Race: Blue Biden Increase Over Obama 1
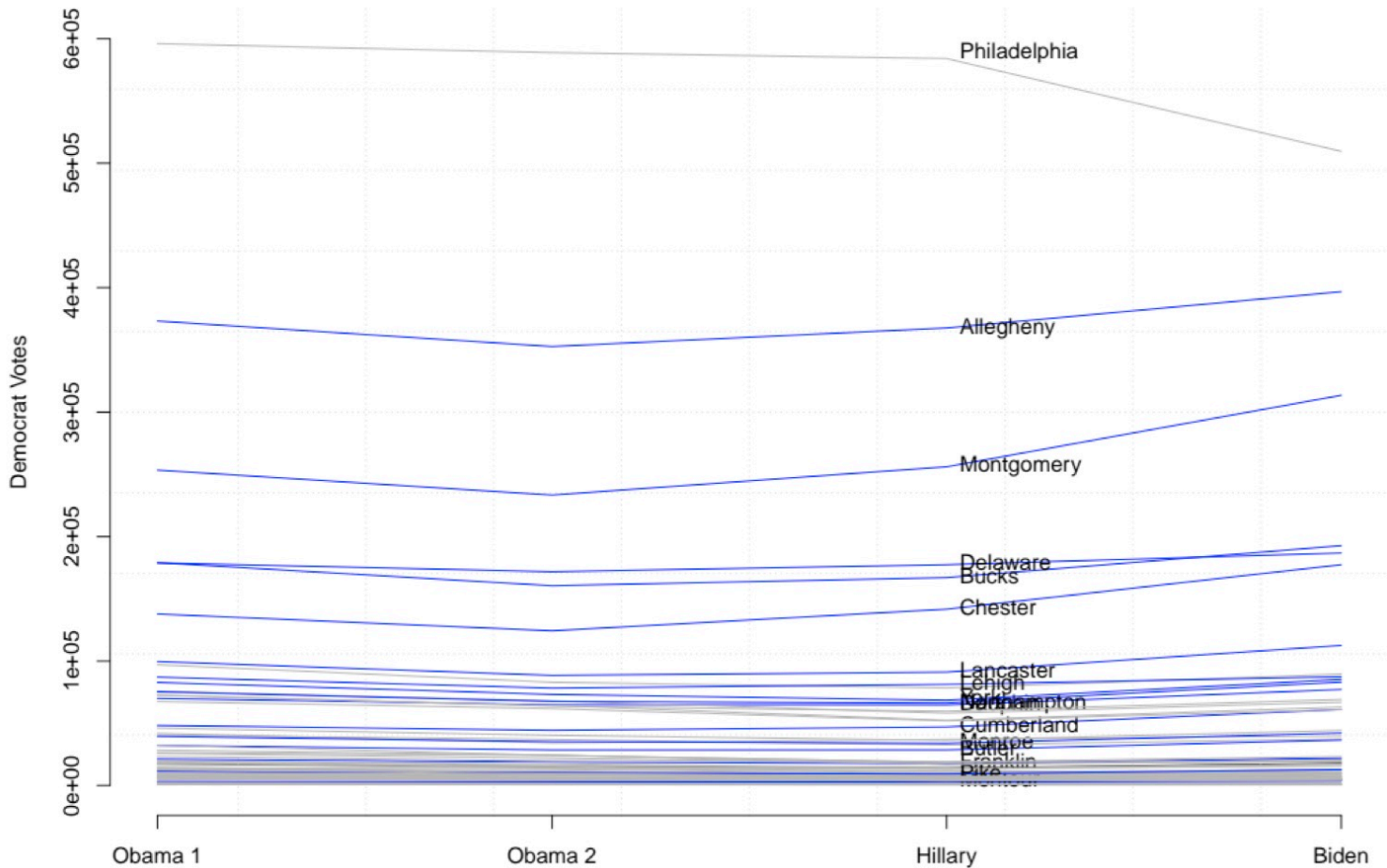
Again, counties which recorded more votes for Biden are highlighted in blue.

Another way to look at this is the total votes cast for any candidate divided by county population (data on population provided by Wikipedia).

Proportion of Total Vote to Population: Blue Biden Over 60%
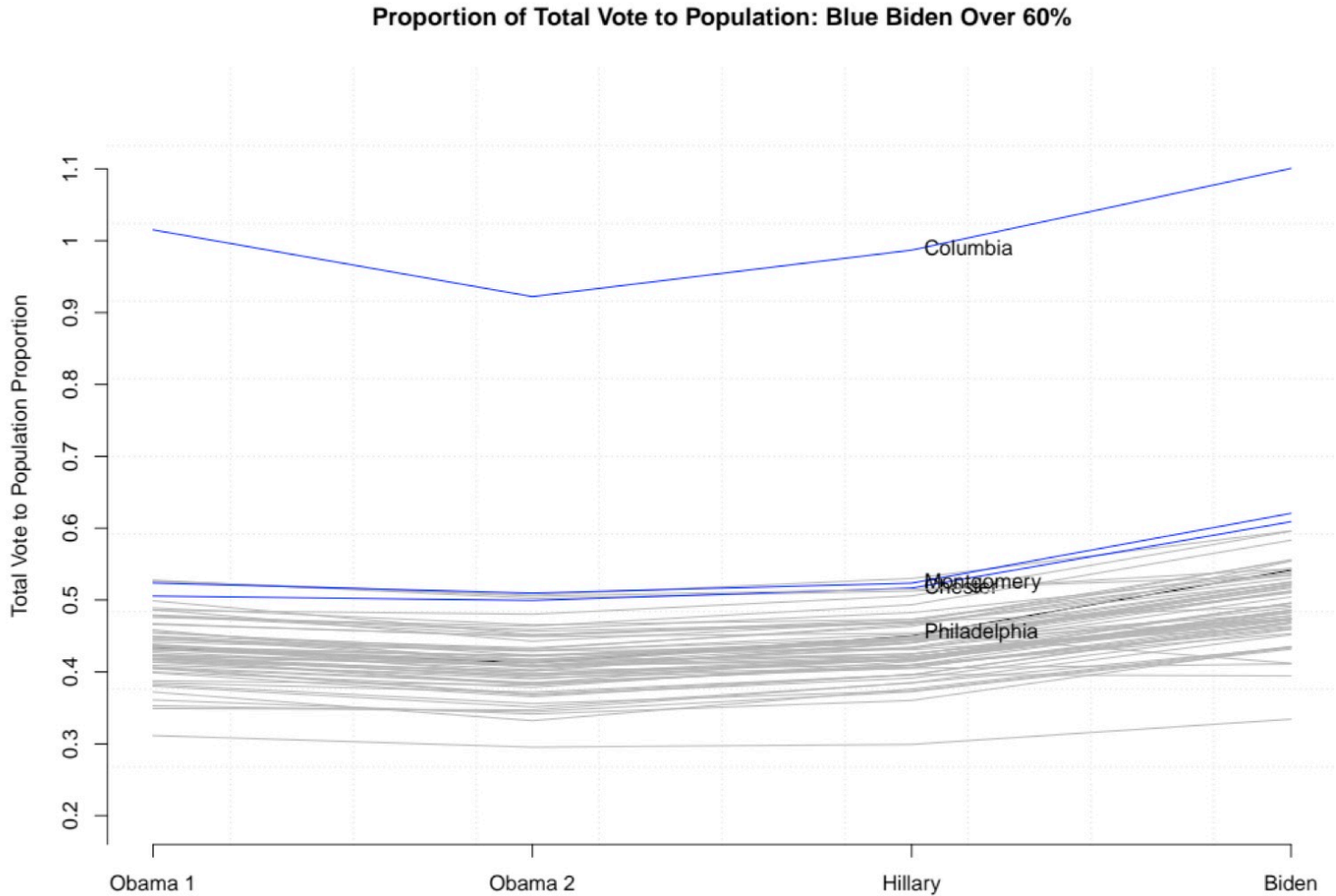
As before, those counties which had higher proportions for Biden than Obama's first run are highlighted in blue. Philadelphia is also noted since it is so large. Columbia appears to have recorded more votes than persons said to live in the county.

The next series of pictures looks at Biden's improvement in total race turnout (votes for all candidates), or not, over his Democrat predecessors' race turnout, by examining the ratio of Biden/Democrat race total votes (for all candidates in any election; this is a measure of turnout) and plotted for each county's proportion of Democrat to Republican registered voters. Counties with proportions < 1 are predominately Republican.

County Level 2020 Biden to Obama First Run Vote Total Ratio by D to R Registration Ratio

Y-axis: 2020 Biden/Obama First Race Vote Totals

X-axis: D/R Registration Ratio

Labels (approximate positions): Pike, Mifflin, Adams, Northampton, Butler, Cumberland, Huntingdon, Chester, Lebanon, Potter, Lancaster, Washington, Montgomery, Lycoming, Blair, Elk, Monroe, Clearfield, Clinton, Tioga, Somerset, Sullivan, Fayette, Bradford, Bucks, Venango, Westmoreland, Wyoming, Dauphin, Susquehanna, Berks, Luzerne, Cameron, Beaver, Lehigh, Clarion, Columbia, Forest, Lawrence, Cambria, Erie, Lackawanna, Allegheny, Mercer, Delaware, Centre, Crawford, Greene, Phila

Greater Turnout
Lesser Turnout
More Rs More Ds

County Level 2020 Biden to Obama Second Run Vote Total Ratio by D to R Registration Ratio

Y-axis: 2020 Biden/Obama Second Race Vote Totals

X-axis: D/R Registrations

Labels (approximate positions): Pike, Adams, Clinton, Mifflin, Wayne, Carbon, Huntingdon, Northampton, Monroe, McKean, Cumberland, Northumberland, Seymour, Washington, Fayette, Potter, Lancaster, Chester, Montgomery, Tioga, Union, Susquehanna, Schuylkill, Luzerne, Clarion, Columbia, Warren, W. Bucks, Westmoreland, Cameron, Dauphin, Beaver, Armstrong, Lehigh, Forest, Erie, Indiana, Centre, Mercer, Allegheny, Greene, Delaware, Crawford, Lackawanna, Phila

Greater Turnout
Lesser Turnout
More Rs More Ds

Another way to look at this is the ratio of Biden votes, i.e. votes just for Biden, over the votes for the other Democrat candidates. This is a measure of popularity, and not turnout per se, like the above figures. Again, this is plotted for each county and by country registration proportion.

County Level Biden to Obama First Run Votess Ratio by D to R Registration Ratio

County Level Biden to Obama Second Run Vote Ratio by D to R Registration Ratio
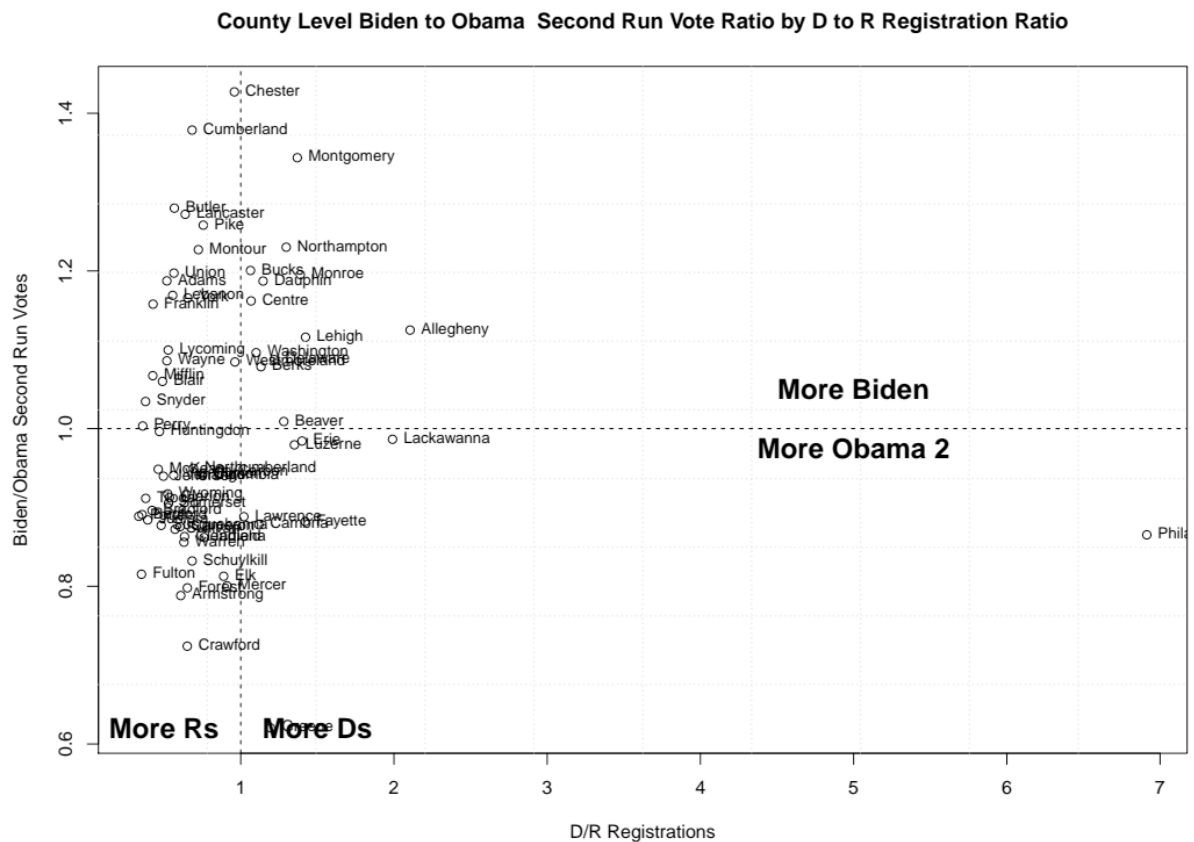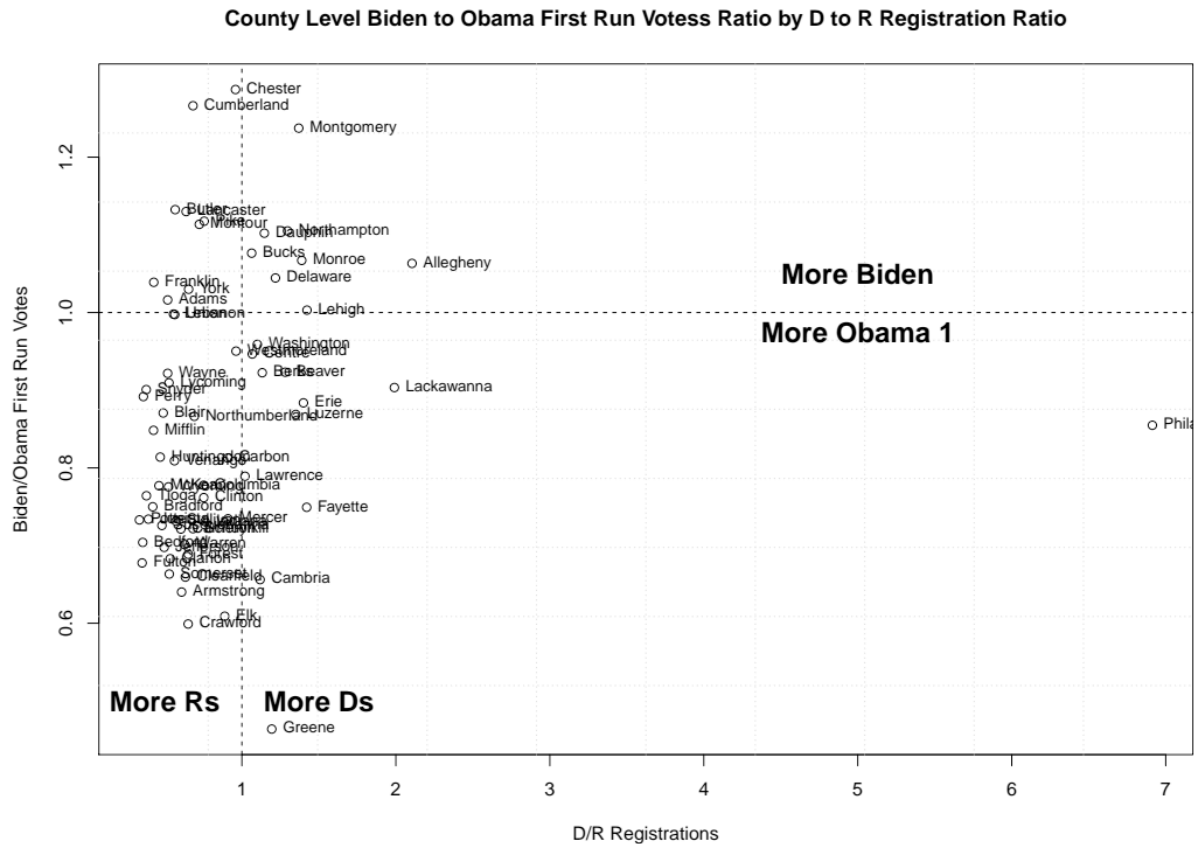
County Level Biden to Hillary Vote Ratio by D to R Registration Ratio
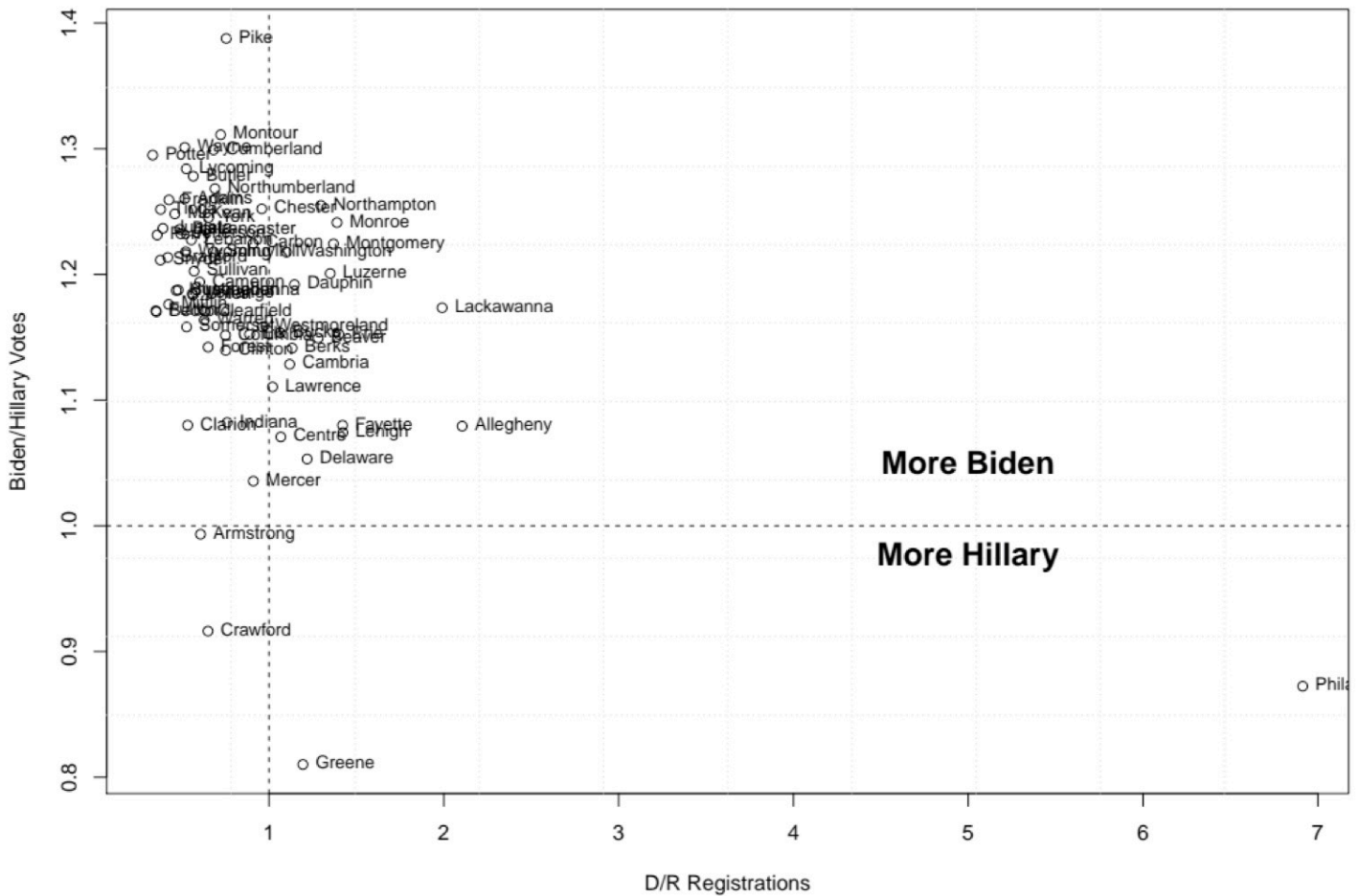
Once more, it's very strange that Biden managed to increase his support over the other Democrat candidates, especially in predominately Republican counties.

Another way to look at this is plotting the proportion of Democrat to Republican registrations by the ratio of Biden to Trump total votes received in the race.

Proportion of Democratic Registrations by Biden/Trump Votes
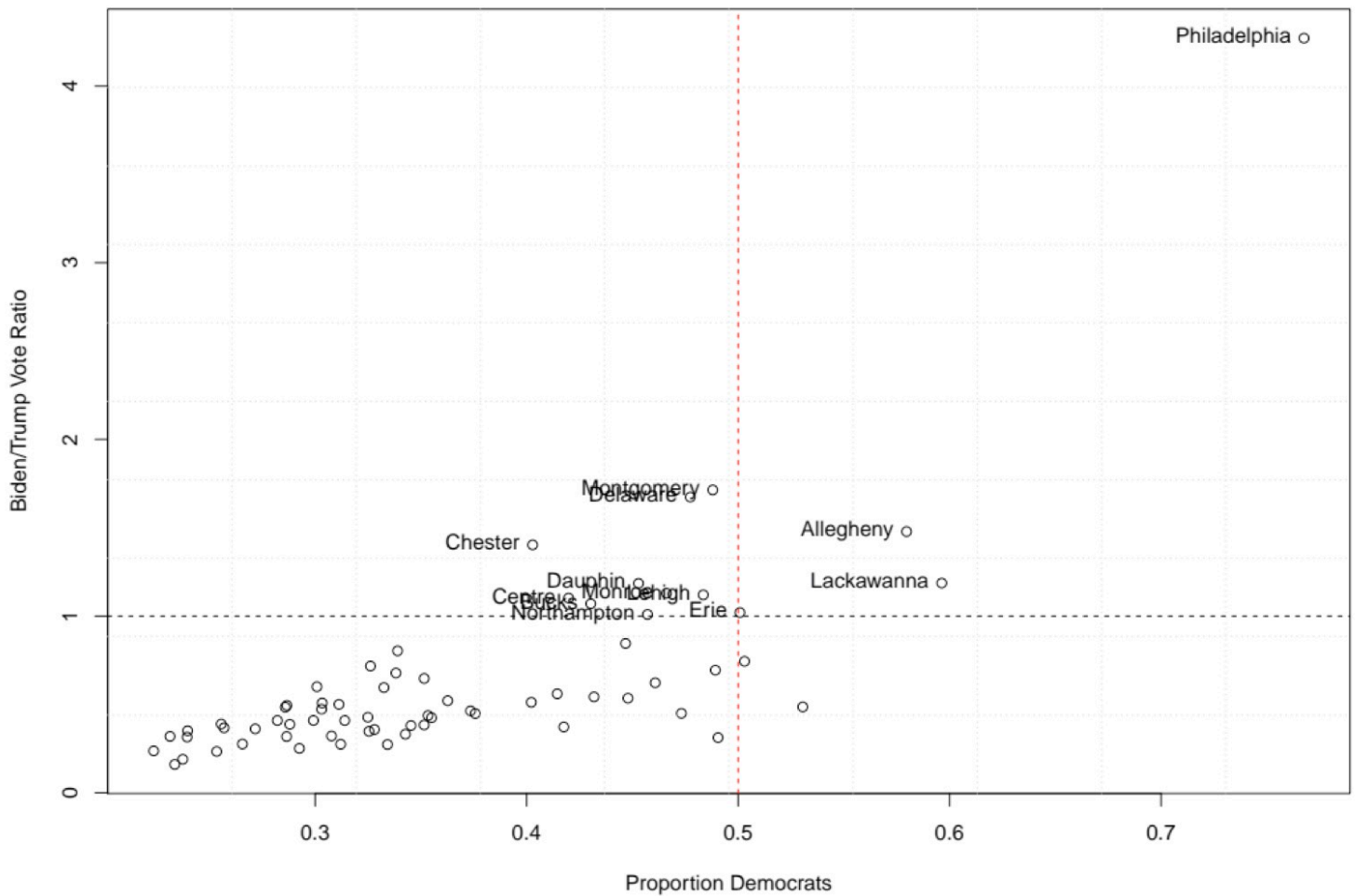
Counties which are predominately Republican have "Proportion Democrats" < 0.5. It's not surprising, necessarily, that Philadelphia county, which is overwhelming Democrat in registrations would have a large Biden/Trump vote ration. But it is very curious several predominately Republican counties would.

County Level 2020 Biden to Hillary Vote Total Ratio by D to R Registration Ratio
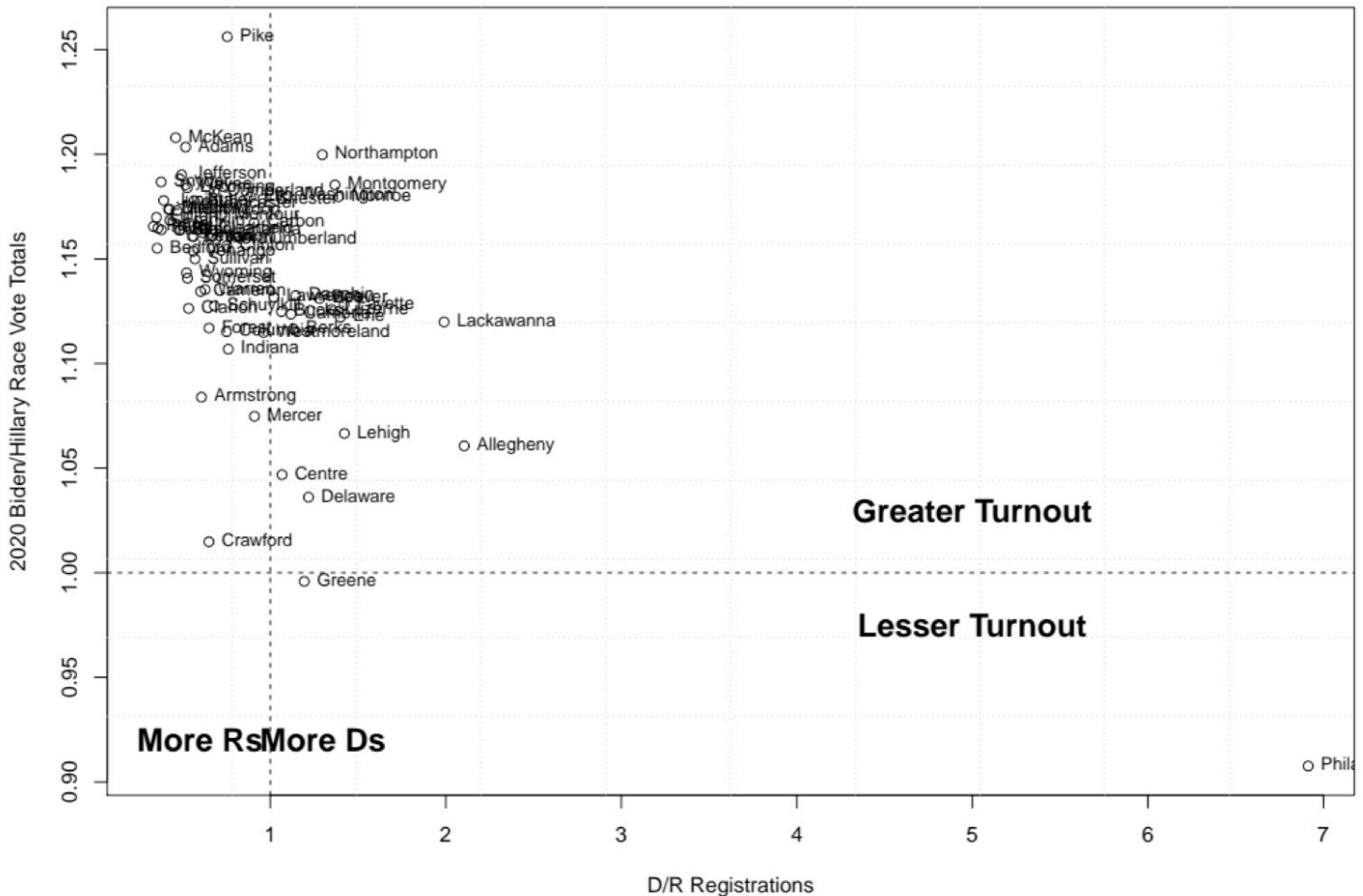

It's very odd the Biden race in total votes bested Obama's first run race total votes (for all candidates) by 20% to 40% in counties which were predominately Republican. In other words, turnout was much higher for 2020 than in Obama's first run against McCain.

MAIL-IN VOTE ANALYSIS

Data on mail-in ballots in Pennsylvania was provided by the same source. It contained the applicant's party affiliation, birth date, the dates the ballots were mailed to applicants, and the dates the ballots were received by authorities. County registration data was used as above, too.

The first thing to note is who requested mail-in ballots. The county ratio of ballots requested by registered voter total is plotted for each party. Dots are red for predominately Republican counties, or Blue for predominately Democrat counties.

Comparing D–R Non–Return Mail–in Ballot Request Rates per Registration County Level
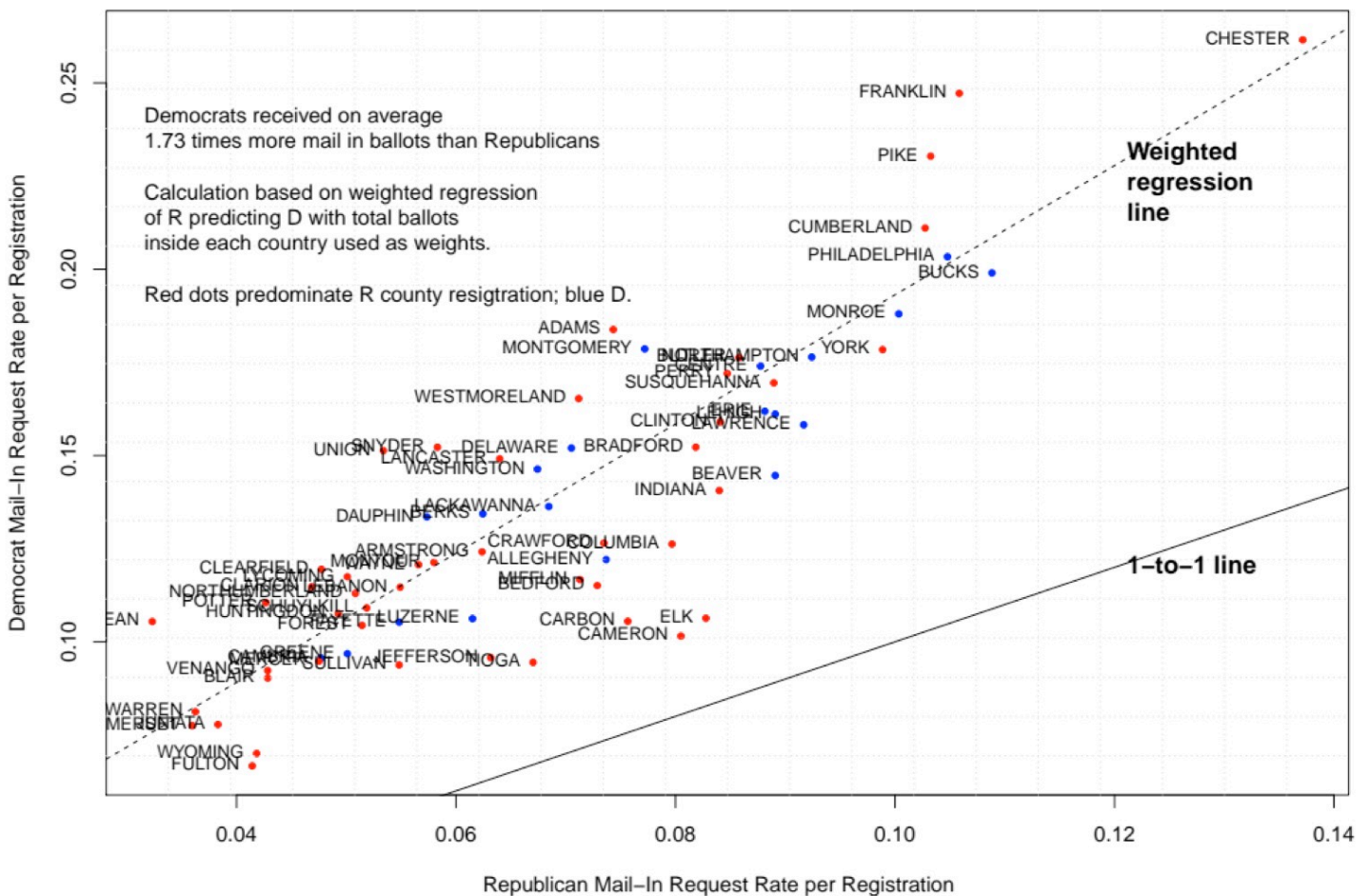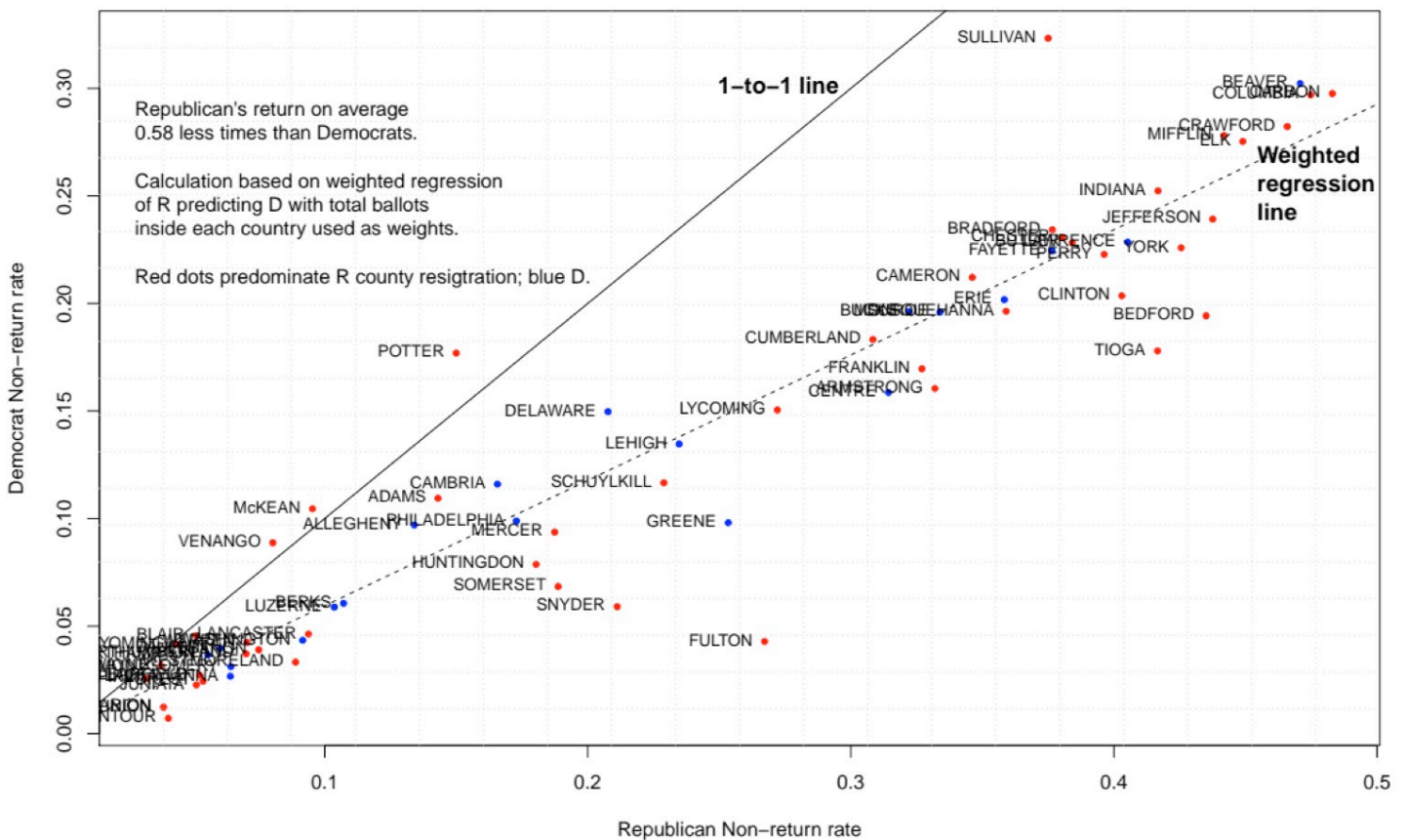
If Republicans and Democrats were recorded as requesting mail-in ballots at equal rates, the counties would line up to the 1-to-1 line. As it is, Democrats were recorded as requesting mail-in ballots 1.73 times as often as Republicans. This was determined by a county-level weighted regression, of Republican ratios predicting Democrat ratios, weighted by the number of mail-in ballots requested in each country (which gives larger counties more weight, as is proper).

For whatever reason, Republicans were recorded as requesting far fewer mail-in ballots than Democrats.

There was a slight difference in mail-in ballots *not* being mailed to Republicans, at 0.9%, versus Democrats, at 0.7%. In other words, proportionally more Republicans than Democrats never had requested mail-in ballots sent to them.

Not every ballot that was mailed out was recorded as returned. Plotted next is the county-level non-return rate for Republicans by Democrats.

Comparing D–R Non–Return Mail–in Ballot Rates County Level

If Republicans were recorded as returning ballots at the same rate as Democrats, counties would line up on the 1-to-1 line. As it is, a weighted regression (as above) shows Republicans were recorded as returning ballots 0.58 times less often than Democrats.

So far we have that just under twice as many Democrats as Republicans were recorded as requesting ballots, and about twice as many Democrats were recorded as returning those ballots. This is also curious and hard to explain logically.

Age did not seem to make any difference in the analysis, nor did breaking any of these charts down by the finer level of State House Districts.

There is one last curiosity. The ballots were mailed so-many days before election day. Ballots mailed to people more days before election day obviously had more time to consider their choices and more time to return their ballots.

Plotted next is the country mean number of days before the elections Republicans were recorded as having ballots mailed out versus Democrats.

Comparing D–R Mean Days Before Election Ballots Mailed to Voters
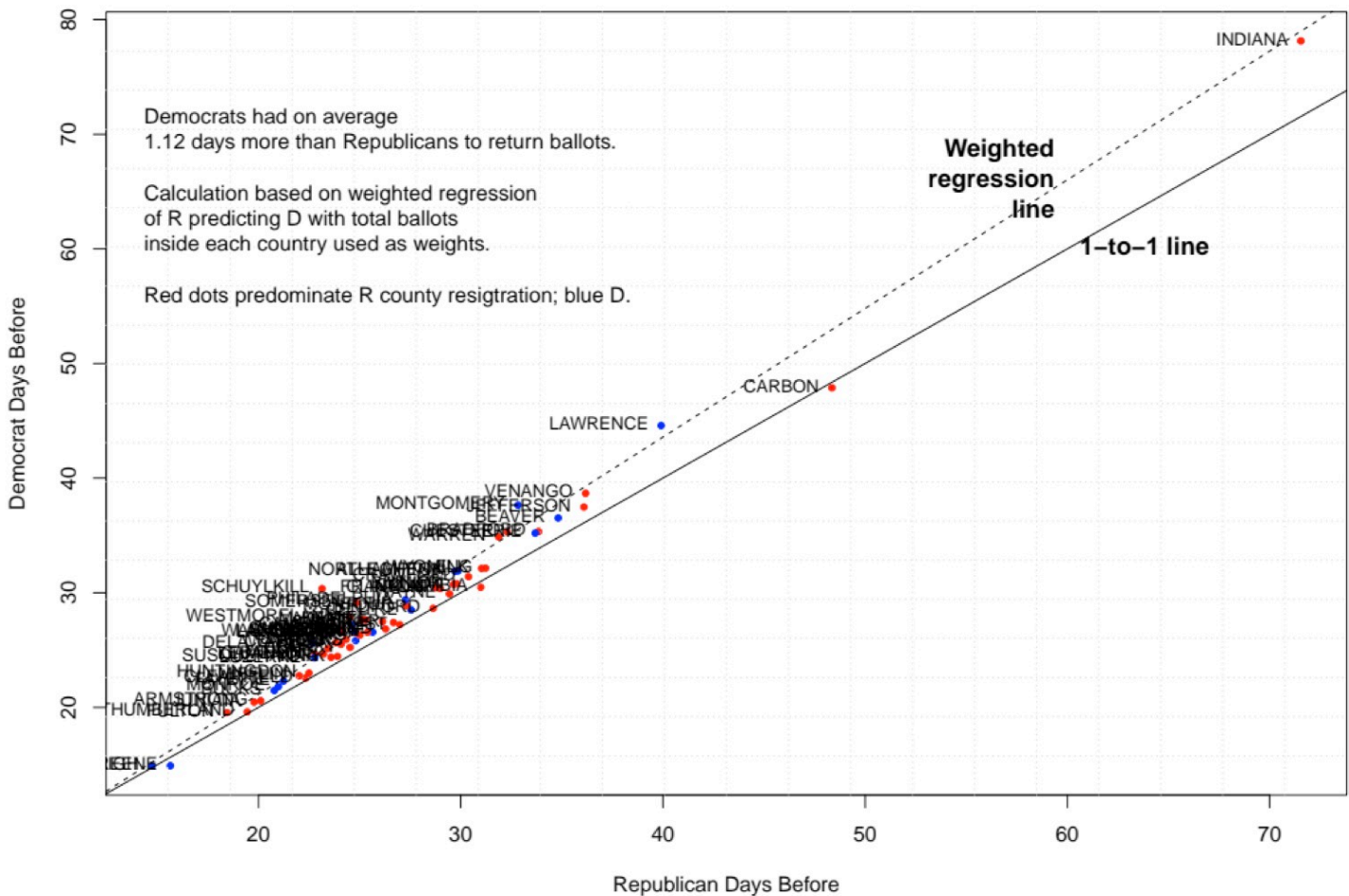
As above, if Republicans were recorded as having as much time as Democrats, the points would fall on the 1-to-1 line. As it is, a weighted regression (as above) showed Democrats had an average 1.12 more days before election than Republicans. Whether or not this is important can be debated, but it was curious to see this happening in almost all counties.

Summary

Several nationally recognized statistical experts were asked to examine some 2020 Pennsylvania voting records, and to identify anything they deemed to be statistically significant anomalies — i.e deviations from the norm.

In the process they by-and-large worked separately, consulted with other experts, analyzed the data they were given from different perspectives, obtained some additional data on their own, etc. — all in a very limited time allotment.

Their one — and only — objective was to try to assure that every legal Pennsylvania vote is counted, and only legal Pennsylvania votes are counted.

The primary takeaway is that ALL of these experts came to the same conclusions:
1) There are some major statistical aberrations in the PA voting records, that are extremely unlikely to occur in a normal (i.e. un-manipulated) setting.
2) The anomalies almost exclusively happened with the Biden votes. Time and again, using a variety of techniques, the Trump votes looked statistically normal.
3) Eleven (out of 67) Pennsylvania counties stood out from all the rest. These counties showed distinctive signs of voting abnormalities — again, all for Biden.
4) The total number of suspicious votes in these counties is 300,000± — which greatly exceeds the reported margin of Biden votes over Trump. (We don't know how many of these are artificial Biden votes, *or* votes switched from Trump to Biden.)
5) These statistical analyses do not prove fraud, but rather provide scientific evidence that the reported results are highly unlikely to be an accurate reflection of how Pennsylvania citizens voted.

As stated in the Executive Summary, our strong recommendation is that (as a minimum):
the five worst of the eleven abnormal PA counties have an immediate audited recount.

If the results of a carefully audited recount are that there is **no** significant change in voting results for all of these five counties (very unlikely), then the authors of this Report recommend that we write off those county deviations as extreme statical flukes, and that the Pennsylvania voting results be certified.

On the other hand, if the results of a carefully audited recount are that there **are** significant changes in voting results for some of these five counties, then the authors of this Report recommend that (as a minimum) that the next six (6) statistically suspicious counties also have an audited recount, prior to any certifying of the Pennsylvania voting results.